Solutions to Chapter 1

1. The intent of this rather vague problem is to get you to compare the two notions, probability as intuition and relative frequency theory. There are many possible answers to how to make the statement "Ralph is probably guilty of theft" have a numerical value in the relative frequency theory. First step is to define a repeatable experiment along with its outcomes. The favorable outcome in this case would be 'guilty.' Repeating this experiment a large number of times would then give the desired probability in a relative frequency sense. We thus see that it may entail a lot of work to attach an objective numerical value to such a subjective statement, if in fact it can be done at all.

   One possible approach would be to look through courthouse statistics for cases similar to Ralph's, similar both in terms of the case itself and the defendant. If we found a sufficiently large number of these cases, ten at least, we could then form the probability $p = n_F/n$, where n_F is the number of favorable (guilty) verdicts, and n is the total number of found cases. Here we effectively assume that the judge and jury are omniscient.

   Another possibility is to find a large number of people with personalities and backgrounds similar to Ralph's, and to expose them to a very similar situation in which theft is possible. The fraction of these people that then steal in relation to the total number of people, would then give an objective meaning to the phrase "Ralph is probably guilty of theft."

2. Note that $D \to 3$, but $3 \not\to D$, i.e., D implies 3 but not the other way around. Thus if we turn over card 2 and find a 3. So what? It was <u>never stated that</u> a $3 \to F$. Likewise, with card 3. On the other hand, if we turn over card 4 and find a D, then the rule is violated. Hence, we must turn over card 4 and card 1, of course.

3. First step here is to decide which kind of probability to use. Since no probabilities are explicitly given, it is reasonable to assume that all numbers are equally likely. Effectively we assume that the wheel is "fair." This then allows us to use the classical theory along with the axiomatic theory to solve this problem. Now we must find the corresponding probability model. We are told in the problem statement that the experiment is "spinning the wheel." We identify the pointed-to numbers as the outcomes ς. The sample space is thus $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$.The total number of outcomes is then 9. The probability of each elemental event $\{i\}$ is then taken as $P[\{i\}] \triangleq p = 1/9$, as in the classical theory. We are also told in the problem statement that the contestant wins if an even number shows. The set of even numbers in Ω is $\{2, 4, 6, 8\}$. We can write this event as a disjoint union of four singleton (atomic) events
$$\{2, 4, 6, 8\} = \{2\} \cup \{4\} \cup \{6\} \cup \{8\}.$$

   Now we can apply axiom 3 of probability to write
$$\begin{aligned} P[\{2, 4, 6, 8\}] &= P[\{2\}] + P[\{4\}] + P[\{6\}] + P[\{8\}] \\ &= \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} \\ &= \frac{4}{9}. \end{aligned}$$

   We have seen that some 'reasonable' assumptions are necessary to transform the given word problem into something that exactly corresponds to a probability model. It turns out that this is a general problem for such word problems, i.e. problems given in natural English.

4. The experiment involves flipping a fair coin 3 times. The outcome of each coin toss is either a head or a tail. Therefore, the sample space of the combined experiment that contains all the possible outcomes of the 3 tosses, is given by

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Since all the coins are fair, all the outcomes of the experiment are equally likely. The probability of each singleton event, i.e. an event with a single outcome, is then $\frac{1}{8}$. We are interested in finding the probability of the event A, which is the event of obtaining 2 heads and 1 tail. There are 3 favorable outcomes for this event given by $A = \{HHT, HTH, THH\}$. Therefore, $P[A] = P[\{HHT\} \cup \{HTH\} \cup \{THH\}] = P[HHT] + P[HTH] + P[THH] = \frac{3}{8}$. Note that we are able to write the probability of the event A as the sum of probability of the singleton events (from Axiom 3) because the singlteon events of any experiment are mutually exclusive. Why?

5. The experiment contains drawing two balls (with replacement) from an urn containing balls numbered 1, 2, and 3. The sample space of the experiment is given by

$$\Omega = \{11, 12, 13, 21, 22, 23, 31, 32, 33\}.$$

The event of drawing a ball twice is said to occur when one of the outcomes 11, 22, or 33 occurs. Therefore, the event of drawing 2 equal balls E, is given by $E = \{11, 22, 33\}$ and $P[E] = P[\{11\}] + P[\{22\}] + P[\{33\}]$. Since the balls are drawn at random, it can assumed that drawing each ball is equally likely. Therefore, the singleton events, or equivalently outcomes of the experiment, are equally likely. Hence, $P[E] = 3(\frac{1}{9}) = \frac{1}{3}$.

6. Let $b_1, b_2, \ldots, b_6$ represent the six balls. Each outcome will be represented by the two balls that were drawn. In the first experiment, the balls are drawn without replacement; hence, the two balls drawn cannot have the same index. Then the sample space containing all the outcomes is given by

$$\begin{aligned}\Omega_1 = \{&b_1b_2, b_1b_3, b_1b_4, b_1b_5, b_1b_6,\\ &b_2b_1, b_2b_3, b_2b_4, b_2b_5, b_2b_6,\\ &b_3b_1, b_3b_2, b_3b_4, b_3b_5, b_3b_6,\\ &b_4b_1, b_4b_2, b_4b_3, b_4b_5, b_4b_6,\\ &b_5b_1, b_5b_2, b_5b_3, b_5b_4, b_5b_6,\\ &b_6b_1, b_6b_2, b_6b_3, b_6b_4, b_6b_5\}.\end{aligned}$$

This can be written compactly as

$$\Omega_1 = \{b_{(i,j)} | 1 \le i \le 6, 1 \le j \le 6, i \neq j\}.$$

If the first ball is replaced before the second draw, then in addition to the outcomes in the earlier part, there are outcomes where both the two balls drawn are the same. The sample space for the new experiment is given by

$$\Omega_2 = \Omega_1 \cup \{b_1b_1, b_2b_2, b_3b_3, b_4b_4, b_5b_5, b_6b_6\}.$$

This can also be written as $\Omega_2 = \{b_{(i,j)} | 1 \le i \le 6, 1 \le j \le 6\}$.

7. Let h_M be the height of the man and h_W be the height of the woman. Each outcome of the experiment can be expressed as a two-tuple (h_M, h_W). Thus

   (a) The sample space Ω is the set of all possible pairs of heights for the man and woman. This is given as
   $$\Omega = \{(h_M, h_W) : h_M > 0, h_W > 0\}.$$

   (b) The event E, which is a subset of Ω is given by
   $$E = \{(h_M, h_W) : h_M > 0, h_W > 0, h_M < h_W\}.$$

8. The word problem describes the physical experiment of drawing numbered balls from an urn. We need to find a corresponding mathematical model. First we form an appropriate event space with meaningful outcomes. Here the physical experiment is 'draw ball from urn,' so the outcome in words is 'particular labeled ball drawn,' which we can identify with its label. So we select as *outcome* in our mathematical model, the number on the drawn ball's face, i.e. the particular label. The outcomes are thus the integers 1,2,3,4,5,6,7,8, 9, and 10. The *sample space* is then $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, and is the set of all ten outcomes. We are told that E is 'the event of drawing a ball numbered no greater than 5.' Thus we define in our event field $E = \{1, 2, 3, 4, 5\}$. The other event specified in the word problem is F 'the event of drawing a ball greater than 3 but less than 9.' In our mathematical event field this corresponds to $F = \{4, 5, 6, 7, 8\}$. Having constructed our sample space with indicated events, we can use elementary set theory to determine the following answers:
   $$\begin{array}{cc} E^c = \{6, 7, 8, 9, 10\}, & F^c = \{1, 2, 3, 9, 10\}, \\ EF = \{4, 5\}, & E \cup F = \{1, 2, 3, 4, 5, 6, 7, 8\}, \\ EF^c = \{1, 2, 3\}, & E^c F = \{6, 7, 8\}, \\ E^c \cup F^c = \{1, 2, 3, 6, 7, 8, 9, 10\}, & \\ (EF^c) \cup (E^c F) = \{1, 2, 3, 6, 7, 8\}, & (EF) \cup (E^c F^c) = \{4, 5, 9, 10\} \\ (E \cup F)^c = \{9, 10\}, & (EF)^c = \{1, 2, 3, 6, 7, 8, 9, 10\}. \end{array}$$

   The last part of the problem asks us to 'express these events in words.' Since we have a mathematical model, we should really more precisely ask what each of these events *corresponds to in words*. We know of course that E corresponds to 'drawing a ball numbered no greater than 5.' We can thus loosely write $E = \{$'drawing a ball numbered no greater than 5'$\}$, although in our mathematical model E is just the set of integers $\{1, 2, 3, 4, 5\}$. So when we write $E =\{$ 'drawing a ball numbered no greater than 5'$\}$, what we really mean is that the event E in our mathematical model corresponds to the physical event 'drawing a ball numbered no greater than 5' mentioned in the word problem. With this caveat in mind, we can then write:
   $$\begin{aligned} E^c &= \{\text{'drawing a ball greater than 5'}\}, \\ F^c &= \{\text{'drawing a ball not in the range 4-8 inclusive'}\}, \\ EF &= \{\text{'drawing a ball greater than 3 and no greater than 5'}\}, \\ & \text{etc.} \end{aligned}$$

9. The sample space containing four equally likely outcomes is given by $\Omega = \{\zeta_1, \zeta_2, \zeta_3, \zeta_4\}$. Two events $A = \{\zeta_1, \zeta_2\}$ and $B = \{\zeta_2, \zeta_3\}$ are given. The required events can be easily obtained by observation.

AB^c = set of outcomes in A and not in $B = \{\zeta_1\}$.
BA^c = set of outcomes in B and not in $A = \{\zeta_3\}$.
AB = set of outcomes in A and $B = \{\zeta_2\}$.
$A \cup B$ = set of outcomes in A or in $B = \{\zeta_1, \zeta_2, \zeta_3\}$.

10. $A = AB \cup AB^c$. This can be proved using the distributive law on

$$A = A\Omega = A(B \cup B^c) = AB \cup AB^c.$$

$A \cup B = (AB^c) \cup (BA^c) \cup (AB)$. Here we first write $A = A(B \cup B^c)$ and $B = B(A \cup A^c)$. Then we can write

$$\begin{aligned} A \cup B &= (A(B \cup B^c)) \cup (B(A \cup A^c)) \\ &= (AB \cup AB^c) \cup (BA \cup BA^c) \\ &= AB \cup AB^c \cup BA \cup BA^c \\ &= AB \cup AB^c \cup BA^c, \end{aligned}$$

using the above laws and formulas. Notice that the above two decompositions are into disjoint sets. From the third axiom of probability, we know that the probability of union of disjoint sets is the sum of the probabilities of the disjoint sets. Therefore, we can add the probabilities over the unions.

11. In a given random experiment there are four *equally likely* outcomes $\zeta_1, \zeta_2, \zeta_3$, and ζ_4. Let the event $A \triangleq \{\zeta_1, \zeta_2\}$.
$P[A] = P[\{\zeta_1, \zeta_2\}] = P[\{\zeta_1\}] + P[\{\zeta_2\}] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. $A^c = \{\zeta_3, \zeta_4\}$,
$P[A^c] = P[\{\zeta_3, \zeta_4\}] = P[\{\zeta_3\}] + P[\{\zeta_4\}] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.
Note that we are told that the four outcomes are equally likely. This means that the four singleton (atomic) events have equal probability. $P[A] = \frac{1}{2} = 1 - P[A^c] = 1 - \frac{1}{2}$.

12. (a) The three axioms of probability are given below

   (a) [label=()]

   (b) For any event A, the probability of the even occuring is always non-negative.

$$P[A] \geq 0.$$

   This ensures that probability is never negative.

   (c) The probability of occurence of the sample space event Ω is one.

$$P[\Omega] = 1.$$

   This ensures that probability of no event exceeds one. The first two axioms ensures that the probability is a quantity between 0 and 1, inclusive.

   (d) For any two events A, B that are disjoint, the probability of the union of the events is the sum of the probabilities of the two events.

$$P[A \cup B] = P[A] + P[B], \text{ when } AB = \phi.$$

   This axiom tells us that the probability of any event can be obtained by the sum disjoint events that constitute the event.

(b) The event $A \cup B$ can be obtained as the disjoint union of the three sets AB, AB^c, A^cB. Hence by applying the third axiom of probability, we obtain

$$\begin{aligned} P[A \cup B] &= P[AB \cup (AB^c \cup A^cB)] \\ &= P[AB] + P[AB^c \cup A^cB] \\ &= P[AB] + P[AB^c] + P[A^cB]. \end{aligned}$$

Now the event A can be written as the disjoint union of AB and AB^c (Axiom 3). Therefore

$$P[A] = P[AB] + P[AB^c] \implies P[AB^c] = P[A] - P[AB]$$

Similarly

$$P[B] = P[AB] + P[A^cB] \implies P[A^cB] = P[B] - P[AB].$$

Therefore $P[A \cup B] = P[AB] + (P[A] - P[AB]) + (P[B] - P[AB]) = P[A] + P[B] - P[AB]$.

13. We first form our mathematical model by setting outcomes $\boldsymbol{\varsigma} = (\varsigma_1, \varsigma_2)$, where ς_1 corresponds to the label on the first ball drawn, and ς_2 corresponds to the label on the second ball drawn. We can also write the outcomes as strings $\boldsymbol{\varsigma} = \varsigma_1\varsigma_2$. The sample space Ω can then be identified with the 2-D array

$$\begin{array}{ccccc} 11 & 12 & 13 & 14 & 15 \\ 21 & 22 & 23 & 24 & 25 \\ 31 & 32 & 33 & 34 & 35 \\ 41 & 42 & 43 & 44 & 45 \\ 51 & 52 & 53 & 54 & 55 \end{array}.$$

There are thus 25 outcomes in the sample space. Now the word problem statement uses the phrase 'at random' to describe the drawing. This is a technical term that can be read 'equally likely.' Thus all the elementary events $\{\varsigma_1\varsigma_2\}$ in our mathematical model must have equal probability, i.e. $P[\{\varsigma_1\varsigma_2\}] = 1/25$. Armed thusly we can attack the given problem as follows. Define the event $E =$\{'sum of labels equals five'\}, or precisely $E = \{41, 32, 23, 14\}$. Then we decompose this event into four singleton events as

$$E = \{41\} \cup \{32\} \cup \{23\} \cup \{14\}.$$

Since different singleton events are disjoint, probability adds, and we have

$$\begin{aligned} P[E] &= \frac{1}{25} + \frac{1}{25} + \frac{1}{25} + \frac{1}{25} \\ &= \frac{4}{25}. \end{aligned}$$

"Dim" ignored that outcome ij is different (distinguishable) from outcome ji. "Dense" talked about the sums and correctly noted that there were nine of them. However, he incorrectly assumed that each sum was equally likely. Looking at our sample space above, we can see that the sum 2 has only one favorable outcome 11, while the sum 6 has five favorable outcomes, just looking at the anti-diagonals of this matrix.

14. First we show $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$.
Let $x \in A \cap (B \cup C)$.
Then $x \in A$ and $x \in (B \cup C)$.
$x \in A$ and $x \in B$ or $x \in C$.

Say if $x \in B$. Then $x \in A$ and $x \in B$ (Step k)
Thus $x \in (A \cap B)$.
And therefore $x \in (A \cap B) \cup (A \cap C)$. Similar arguments can be made if we consider $x \in C$ in step k, in which case we will show that $x \in (A \cap C)$ and hence $x \in (A \cap B) \cup (A \cap C)$.
Thus we have shown that $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$.

Now we show that $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$.
Suppose $x \in (A \cap B) \cup (A \cap C)$. Then $x \in (A \cap B)$ or $x \in (A \cap C)$.
Say $x \in (A \cap B)$
Then $x \in A$ and $x \in B$.
Or $x \in A$ and $x \in (B \cup C)$. Or in other words, $x \in A \cap (B \cup C)$.

Similar arguments can be used to show that if $x \in (A \cap C)$, then $x \in A \cap (B \cup C)$.
Thus $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$.

Thus we have shown that both sets are contained in each other. Hence $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

15. We use the set identity $\Omega = A \cup A^c$. Since this union is disjoint, by the additivity of probability (i.e. axiom 3), we get $1 = P[\Omega] = P[A] + P[A^c]$, which with rearranging becomes the desired result.

16. (a) $A \cap C = \{1, 2\} \cap \{4, 5, 6\} = \phi$. Therefore,

$$
\begin{aligned}
P[A \cap C] &= P[\phi] \\
&= 1 - P[\Omega] \\
&\quad (\because \Omega \cap \Phi = \phi, 1 = P[\Omega \cup \phi] = P[\Omega] + P[\phi]) \\
&= 1 - 1 \ \text{(because } P[\Omega] = 1) \\
&= 0.
\end{aligned}
$$

(b) $P[A \cup B \cup C] = P[\{1, 2\} \cup \{2, 3\} \cup \{4, 5, 6\} = P[\{1, 2, 3, 4, 5, 6\}] = P[\Omega] = 1$.

(c) We see that $B \cap C = \phi$ and so $P[BC] = 0$. For B and C to be independent, $P[BC] = P[B]P[C]$. Therefore, if either $P[B] = 0$ or $P[C] = 0$ or both are zeros, B and C will be independent.

17. This problem uses only set theory and just two axioms of probability to get these general results.

(a) We need to show $P[\phi] = 0$. We write the disjoint decomposition $\Omega = \Omega \cup \phi$ and then use the additivity of probability (axiom 3) to get

$$
\begin{aligned}
P[\Omega] &= P[\Omega \cup \phi] \\
&= P[\Omega] + P[\phi].
\end{aligned}
$$

So we must have $P[\phi] = 0$.

(b) Using set theory, we can write the disjoint decomposition

$$E = EF^c \cup EF.$$

Then by axiom 3, the additivity of probability, we have

$$
\begin{aligned}
P[E] &= P[EF^c \cup EF] \\
&= P[EF^c] + P[EF],
\end{aligned}
$$

or what is the same $P[EF^c] = P[E] - P[EF]$.

(c) Here we simply note $E \cup E^c = \Omega$ is a disjoint decomposition, so that again by axiom 3,

$$\begin{aligned} P[\Omega] &= P[E] + P[E^c] \\ &= 1, \quad \text{by axiom 2,} \end{aligned}$$

which is the same as $P[E] = 1 - P[E^c]$.

18. The outcome is the result of a probabilistic experiment. An event is a collection (set) of outcomes. The field of events is the complete collection of events that are relevant for the given probability problem.

19. We start with the mutually exclusive decomposition

$$A \cup B = AB^c \cup AB \cup A^cB,$$

yielding $P[A \cup B] = P[AB^c] + P[AB] + P[A^cB]$. Then consider the two simple disjoint decompositions

$$A = AB^c \cup AB \quad \text{and} \quad B = A^cB \cup AB,$$

which yield $P[A] = P[AB^c] + P[AB]$ and $P[B] = P[A^cB] + P[AB]$. Putting them all together, we have

$$\begin{aligned} P[A \cup B] &= P[AB^c] + P[AB] + P[A^cB] \\ &= (P[A] - P[AB]) + P[AB] + (P[B] - P[AB]) \\ &= P[A] + P[B] - P[AB]. \end{aligned}$$

20. From Eq. 1.4-3, we see that $E \oplus F = (E - F) \cup (F - E) = EF^c \cup E^cF$. We see that EF^c and E^cF are disjoint, i.e., $(EF^c) \cap (E^cF) = \phi$. Therefore, the probability of the union of EF^c and E^cF are the sum of the probabilities of the two events. In other words,

$$P(E \oplus F) = P(EF^c \cup E^cF) = P(EF^c) + P(E^cF).$$

21. We have already (Problem 17) seen that we can write $P[EF^c] = P[E] - P[EF]$ and $P[E^cF] = P[F] - P[EF]$. Therefore, $P(E \oplus F) = P(EF^c) + P(E^cF) = P[E] + P[F] - 2P[EF]$.

22. (a) For simplicity associate as follows: cat=1, dog=2, goat=3, and pig=4. The outcomes ξ then become the integers 1,2,3, and 4. The sample space $\Omega = \{1, 2, 3, 4\}$. For probability information we are given:

$$P[\{1, 2\}] = 0.9, P[\{3, 4\}] = 0.1, P[\{4\}] = 0.05, \text{ and } P[\{2\}] = 0.5.$$

Now for every event in our field of events, we must be able to specify the probability. This is equivalent to being able to supply the probability for all the singleton events. To see if we can do this, we note that singleton events $\{1\}$ and $\{3\}$ are missing probabilities, so we first write

$$\begin{aligned} \{1\} &= \{1, 2\} - \{2\}, \text{ so that} \\ P[\{1\}] &= P[\{1, 2\}] - P[\{2\}] = 0.9 - 0.5 = 0.4. \end{aligned}$$

Doing the same for the other missing singleton probability $P[\{3\}]$, we write

$$\begin{aligned} \{3\} &= \{3,4\} - \{4\}, \text{ so that} \\ P[\{3\}] &= P[\{3,4\}] - P[\{4\}] = 0.1 - 0.05 = 0.05. \end{aligned}$$

Thus we have enough probability information for all the singleton events, and hence all $16 = 2^4$ subsets of $\Omega = \{1,2,3,4\}$. The appropriate field $\mathcal{F}$ of events then consists of the following events along with their probabilities:

$$\begin{array}{cc} \{1\}, & P[\{1\}] = 0.4, \\ \{2\}, & P[\{2\}] = 0.5, \\ \{3\}, & P[\{3\}] = 0.05, \\ \{4\}, & P[\{4\}] = 0.05, \\ \{1,2\}, & P[\{1,2\}] = 0.9, \\ \{1,3\}, & P[\{1,3\}] = 0.45, \\ \{1,4\}, & P[\{1,4\}] = 0.45, \\ \{2,3\}, & P[\{2,3\}] = 0.55, \\ \{2,4\}, & P[\{2,4\}] = 0.55, \\ \{3,4\}, & P[\{3,4\}] = 0.1, \\ \{1,2,3\}, & P[\{1,2,3\}] = 0.95, \\ \{1,2,4\}, & P\{1,2,4\}] = 0.95, \\ \{1,3,4\}, & P[\{1,3,4\}] = 0.5, \\ \{2,3,4\}, & P[\{2,3,4\}] = 0.6, \\ \{1,2,3,4\}(=\Omega), & P[\{1,2,3,4\}] = 1 = P[\Omega], \\ \phi, & P[\phi] = 0. \end{array}$$

(b) Now the above is not an appropriate field of events if some of the events do not have known probabilities. So if $P[\text{'pig'} = \{4\}] = 0.05$ is removed, then we cannot determine the probabilities of some of the above events. In particular we cannot find $P[\{3\}]$. The alternative then is to treat $\{3,4\}$, whose probability is still given, as a singleton and form a smaller field with just the 8 events formed by unions of $\{1\}$, $\{2\}$, and $\{3{,}4\}$. The resulting field, along with its probabilities is as follows:

$$\begin{array}{cc} \{1\}, & P[\{1\}] = 0.4, \\ \{2\}, & P[\{2\}] = 0.5, \\ \{1,2\}, & P[\{1,2\}] = 0.9, \\ \{3,4\}, & P[\{3,4\}] = 0.1, \\ \{1,3,4\}, & P[\{1,3,4\}] = 0.5, \\ \{2,3,4\}, & P[\{2,3,4\}] = 0.6, \\ \{1,2,3,4\}(=\Omega), & P[\{1,2,3,4\}] = 1 = P[\Omega], \\ \phi, & P[\phi] = 0. \end{array}$$

23. First we show that $A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C)$.
Suppose $x \in A \cap (B \cup C)$
Then $x \in A$
Therefore $x \in (A \cup B)$, and $x \in (A \cup C)$
Hence, $x \in (A \cup B) \cap (A \cup C)$.

Now we show that $(A \cup B) \cap (A \cup C) \subset A \cup (B \cap C)$.
Suppose $x \in (A \cup B) \cap (A \cup C)$

Then $x \in (A \cup B)$ and $x \in (A \cup C)$
$x \in A or B$ and $x \in A or C$
If $x \in A$, then $x \in A \cup (B \cap C)$ (because $A \subset (A \cup (B \cap C))$)
If $x \in A$, then $x \in B$ and $x \in C$.
Or in other words, $x \in (B \cap C)$
$x \in A \cup (B \cap C)$.

Thus we have shown that both the sets are contained in each other. Therefore, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

24. The probability of A is $P[A] = P[\{\zeta_1, \zeta_2\}] = P[\{\zeta_1\}] + P[\{\zeta_2\}] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. The event (set) A^c in terms of the outcomes is $A^c = \{\zeta_3, \zeta_4\}$. The probabilty of A^c is $P[A^c] = P[\{\zeta_3, \zeta_4\}] = P[\{\zeta_3\}] + P[\{\zeta_4\}] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. Note that we are told that the four outcomes are equally likely. This means that the four singleton (atomic) events have equal probability. We verify $P[A] = \frac{1}{2} = 1 - P[A^c] = 1 - \frac{1}{2}$.

25. The composition of the urn is: (a), (a), (b), (b), (ab), (ab), (ab), (ab). $P[A] = 6/8$, $P[B] = 6/8$, $P[AB] = n_{ab}/n_T = 4/8$ is not equal to $P[A]P[B] = 9/16$. Therefore A and B are not independent.

26. Let $n_i, i = 1, 2$ represent the outcome of the ith toss. Since the tosses are independent:

$$P[n_1, n_2] = P[n_1]P[n_2] = \frac{1}{6} \cdot \frac{1}{6}$$

$$\begin{aligned} P[n_1 + n_2 = 7 | n_1 = 3] &= P[n_2 = 4 | n_1 = 3] \\ &= \frac{P[n_1 = 3, n_2 = 4]}{P[n_1 = 3]} \\ &= \frac{P[n_1 = 3]P[n_2 = 4]}{P[n_1 = 3]} \\ & \text{(because tosses are independent)} \\ &= \frac{1}{6} \end{aligned}$$

27. Clearly

$$P[A] = \frac{4}{52} \quad \text{and} \quad P[B] = \frac{26}{52} = \frac{1}{2}.$$

Then $P[AB] = P[\{\text{pick one of two red aces in 52 cards}\}] = \frac{2}{52}$. Is $P[AB] = P[A]P[B]$? Now

$$\begin{aligned} P[AB] = \frac{2}{52} &= \frac{4}{52}\frac{1}{2} \\ &= P[A]P[B], \end{aligned}$$

so, yes A and B are independent events.

28. Since it is a fair die, the successive tosses are independent with probability $p = 1/6$ for each face. From the provided information, we equivalently want the probability of getting a total of 5 on the two remaining tosses. This can happen in just 4 equally likely outcomes, i.e. (4,1), (3,2), (2,3), and (1,4). The desired probability this then $4/36 = 1/9$.

29. We can look at the compound outcomes $\boldsymbol{\varsigma} = (\varsigma_1, \varsigma_2)$ as corresponding to the locations in the $9{\times}9$ array

$$\begin{array}{cccccccc}
11 & 21 & 31 & 41 & 51 & 61 & \cdots & 91 \\
12 & 22 & 32 & 42 & 52 & \cdots & & \vdots \\
13 & 23 & 33 & 43 & \cdots & & & \\
14 & 24 & 34 & \cdots & & & & \\
15 & 25 & \cdots & & \ddots & & & \\
16 & \cdots & & & & & & \\
\vdots & & & & & & & \\
 & & & & & & \ddots & \vdots \\
19 & \cdots & & & & & \cdots & 99
\end{array},$$

with 81 equally likely outcomes. We agree to call the sample space for the first experiment Ω_1, the sample space for the second experiment Ω_2, and the compound sample space simply Ω. To get the sum $\Sigma \triangleq N_1 + N_2 = 7$, we need one of the following outcomes

$$16, 25, 34, 43, 52, 61, \quad \text{located on a } 45^\circ \text{ diagonal in the above table.}$$

So there are 6 favorable outcomes for the event $\{\Sigma = 7\}$. The event $\{\Sigma = \text{odd}\}$ contains 40 outcomes and the event $\{\Sigma = \text{even}\}$ contains the remaining $81 - 40 = 41$ even-sum outcomes. Now the joint event $\{\Sigma = 7\} \cap \{\Sigma = \text{odd}\} = \{\Sigma = 7\}$ since the sum 7 is an odd number. We can now calculate the needed probabilities

$$P[\{\Sigma = \text{odd}\}] = \frac{40}{81} \quad \text{and} \quad P[\{\Sigma = 7\}] = \frac{6}{81}.$$

The answer for the first question is then

$$\begin{aligned}
P[\{\Sigma &= 7\}|\{\Sigma = \text{odd}\}] = \frac{P[\{\Sigma = 7\} \cap \{\Sigma = \text{odd}\}]}{P[\{\Sigma = \text{odd}\}]}, \quad \text{(by definition)} \\
&= P[\{\Sigma = 7\}]/P[\{\Sigma = \text{odd}\}], \qquad \text{(by above result)} \\
&= 6/40.
\end{aligned}$$

The next question is to find $P[(\{N_1 > 7\} \times \Omega_2) \cup (\Omega_1 \times \{N_2 > 7\})|\{\Sigma > 10\}]$. For simplicity of notation, let's agree to write the compound events $\{N_1 > 7\} \times \Omega_2$ and $\Omega_1 \times \{N_2 > 7\}$ as simply $\{N_1 > 7\}$ and $\{N_2 > 7\}$, respectively, for the rest of this calculation. So we must count the relevant number of outcomes from the above 9×9 array, where the various sums are found on 45° diagonals. For the event $\{\Sigma > 10\}$, we count 36 outcomes. For the joint event $(\{N_1 > 7\} \cup \{N_2 > 7\}) \cap \{\Sigma > 10\}$, we find it easier to consider the set of outcomes that make up the remainder of the event $\{\Sigma > 10\}$, i.e. the event $\{N_1 \leq 7\} \cap \{N_2 \leq 7\} \cap \{\Sigma > 10\}$ which is equal, in words, to the event '$N_1 \leq 7$ and $N_2 \leq 7$ and $\Sigma > 10$ '. We could call this the *complement with respect to* $\{\Sigma > 10\}$ of the event $(\{N_1 > 7\} \cup \{N_2 > 7\}) \cap \{\Sigma > 10\}$. Anyway, we find from the 9×9 array that the number of outcomes in $\{N_1 \leq 7\} \cap \{N_2 \leq 7\}) \cap \{\Sigma > 10\}$ is composed of the following 10 cases: $\sum = 11 = 6 + 5 = 5 + 6 = 7 + 4 = 4 + 7$ and $\sum = 12 = 5 + 7 = 7 + 5 = 6 + 6$ and $\sum = 13 = 6 + 7 = 7 + 6$ and $\sum = 14 = 7 + 7$. So we subtract these 10 outcomes from the 36 outcomes in the event $\{\Sigma > 10\}$ to obtain 26 outcomes in the compound event $(\{N_1 > 7\} \cup \{N_2 > 7\}) \cap \{\Sigma > 10\}$. The relevant probabilities are then

$$P[\{\Sigma > 10\}] = \frac{36}{81} \quad \text{and} \quad P[(\{N_1 > 7\} \cup \{N_2 > 7\}) \cap \{\Sigma > 10\}] = \frac{26}{81}.$$

The desired conditional probability is then

$$P[(\{N_1 > 7\} \cup \{N_2 > 7\})|\{\Sigma > 10\}] = \frac{26/81}{36/81} = \frac{26}{36} \approx 0.72.$$

Finally to compute $P[\{\Sigma = \text{odd}\}|\{N_1 > 8\}]$, we proceed as follows. For the combined experiment, we know there is only one possibility for $N_1 > 8$ and that is $N_1 = 9$, along with any value for N_2. Thus there are 9 outcomes in the compound event $\{N_1 > 8\}$[1], so that it's probability is 9/81. Now the joint event $\{\Sigma = \text{odd}\} \cap \{N_1 > 8\} = \{N_1 = 9\} \cap \{\Sigma = \text{odd}\} = \{(9,2),(9,4),(9,6),(9,8)\}$ with four outcomes. Thus since all outcomes are equally likely, we have

$$P[\{\Sigma = \text{odd}\} \cap \{N_1 > 8\}] = \frac{4}{81}.$$

The desired conditional probability is then

$$\begin{aligned} P[\{\Sigma &= \text{odd}\}|\{N_1 > 8\}] = \frac{P[\{\Sigma = \text{odd}\} \cap \{N_1 > 8\}]}{P[\{N_1 > 8\}]} \\ &= \frac{4/81}{9/81} = \frac{4}{9} \approx 0.44. \end{aligned}$$

30. We are given that $P[D] = 0.001$, where D is the event 'disease is present.'. Let T denote the event 'test is positive,' so that T^c is the event 'test is negative.' We are additionally given $P[T|D] = 1$ and $P[T|D^c] = 0.005$. We are asked to compute $P[D|T]$, i.e. the probability that 'disease is present given the test is positive.' We use Bayes' rule and Theorem as follows

$$\begin{aligned} P[D|T] &= \frac{P[DT]}{P[T]} \\ &= \frac{P[T|D]P[D]}{P[T|D]P[D] + P[T|D^c]P[D^c]} \\ &= \frac{1 \times 0.001}{1 \times 0.001 + 0.005 \times 0.999} \\ &= \frac{1}{1 + 4.995} \approx 0.167. \end{aligned}$$

Thus in only about 17% of the cases will a positive test result actually confirm that you suffer from the disease. The other 83% of the time you will be needlessly worried!

31. Let S_1 denote the set of occupations and let S_2 denote the set of interests and/or hobbies. Then

$$\begin{aligned} S_1 &= \{\text{'office manager', 'engineer', 'doctor', 'teacher', ...}\}, \\ S_2 &= \{\text{'nat. defense', 'books', 'music', 'cooking',...}\}. \end{aligned}$$

Let X denote Henrietta's occupation and Y her interests. Then

$$\begin{aligned} &P[X = \text{'office manager'}, Y = \text{'nat. defense'}] \\ &\quad = P[X = \text{'office manager'}]P[Y = \text{'nat. defense'}|X = \text{'office m} \\ &\quad \leq P[X = \text{'office manager'}], \end{aligned}$$

since $0 \leq P[Y = \text{'nat. defense'}|X = \text{'office manager'}] \leq 1$.

[1] Remember, we decided above to write simply $\{N_1 > k\}$ for the compound event $\{N_1 > k\} \times \Omega_2$. This since, in this problem, we only compute probabilities for events in the compound experiment.

32. Directly from the problem statement

$$\begin{aligned} P[X=3] &= 3 \cdot P[X=1], \\ P[X=2] &= 2 \cdot P[X=1]. \end{aligned}$$

But we also know $P[X=3]+P[X=2]+P[X=1]=1$ which is always true by axiom 2 $P[\Omega]=1$. Therefore $P[X=1]=1/6$, $P[X=2]=1/3$, and $P[X=3]=1/2$. Using Bayes' Theorem, we then compute

$$\begin{aligned} P[X=1|Y=1] &= \frac{P[Y=1|X=1]P[X=1]}{\sum_{i=1}^{3} P[Y=1|X=i]P[X=i]} \\ &= \frac{(1-\alpha)1/6}{(1-\alpha)\frac{1}{6}+\frac{\beta}{2}\frac{1}{3}+\frac{\gamma}{2}\frac{1}{2}} \\ &= \frac{1-\alpha}{1-\alpha+\beta+\frac{3}{2}\gamma}. \end{aligned}$$

33. Let

$$\begin{aligned} A &\triangleq \{\text{examinee knows}\}, \\ B &\triangleq \{\text{examinee guesses}\}, \text{ and} \\ C &\triangleq \{\text{getting right answer}\}\ . \end{aligned}$$

Then $P[A]=p, P[B]=1-p, P[C|A]=1$, and $P[C|B]=1/m$. So

$$\begin{aligned} P[A|C] &= \frac{P[C|A]P[A]}{P[C]} \\ &= \frac{1 \cdot p}{P[C|A]P[A]+P[C|B]P[B]} \\ &= \frac{p}{p+\frac{1}{m}(1-p)} \\ &= \frac{mp}{mp+(1-p)}. \end{aligned}$$

34. There are N contestants and only one most beautiful. Hence

$$P[\{\text{pick most beautiful}\}] = 1/N.$$

35. Let

$$\begin{aligned} \widetilde{A} &\triangleq \{\text{random drawn chip} \in A\}, \\ \widetilde{B} &\triangleq \{\text{random drawn chip} \in B\}, \text{ and} \\ \widetilde{C} &\triangleq \{\text{random drawn chip} \in C\}. \end{aligned}$$

Also, let $D \triangleq \{\text{random drawn chip is defective}\}$. Then

$$\begin{aligned} P[D] &= P[D|\widetilde{A}]P[\widetilde{A}]+P[D|\widetilde{B}]P[\widetilde{B}]+P[D|\widetilde{C}]P[\widetilde{C}] \\ &= 0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.40 \\ &= 0.0345. \end{aligned}$$

Hence

$$\begin{aligned}
P[\widetilde{A}|D] &= \frac{P[D|\widetilde{A}]P[\widetilde{A}]}{P[D]} = \frac{0.05 \times 0.25}{0.0345} \doteq 0.363 \\
P[\widetilde{B}|D] &= \frac{P[D|\widetilde{B}]P[\widetilde{B}]}{P[D]} = \frac{0.04 \times 0.35}{0.0345} \doteq 0.406 \\
P[\widetilde{C}|D] &= \frac{P[D|\widetilde{C}]P[\widetilde{C}]}{P[D]} = \frac{0.02 \times 0.40}{0.0345} \doteq 0.232
\end{aligned}$$

36. From the example

$$P[C] \simeq \frac{k}{N} \log \frac{N}{k}$$

We set $x \triangleq k/N$ and construct the following table.

x	$P[C]$	x	$P[C]$
0.0	0.0	0.5	0.346
0.1	0.23	0.6	0.31
0.2	0.32	0.7	0.25
0.3	0.361	0.8	0.18
0.4	0.367	0.9	0.10

The peak is quite shallow, therefore the choice of k is not critical near the peak.

37. (a) If we associate the 103 villagers with $r = 103$ balls and the $n = 30$ tents with 30 cells, this becomes a classical occupancy problem.

(b) The result is given by Eq.1.8-6, which is repeated here as

$$\begin{aligned}
\binom{n+r-1}{r} &= \binom{30+103-1}{103} \\
&= \frac{132!}{103!29!}.
\end{aligned}$$

(c) The result is

$$\begin{aligned}
\binom{r-1}{r-n} &= \binom{103-1}{103-30} \\
&= \frac{102!}{73!29!}.
\end{aligned}$$

To obtain numerical evaluations of these factorial expressions, one might want to use Stirling's formulas:

$$n! \approx (2\pi)^{1/2} n^{n+1/2} e^{-n}.$$

38. The most natural set of outcomes here are the strings (or vectors) of length r, indicating where each ball has landed. There are n^r such strings. They are all equally likely. The number of favorable outcomes would be $r!$ since there are r choices for the first preselected location, $r-1$ choices for the second location, etc. The desired probability is then $P = r!/n^r$. Now, since the balls are indistinguishable, we could have considered the so-called distinguishable outcomes, $\binom{n+r-1}{r}$ in number, however from the description of the experiment in the problem statement, they would not all be equally likely. So we could not rely on classical theory then to give us the probabilities of these outcomes.

39. As in problem 1.38, the number of favorable ways is $r!$. However, the total number of ways is *not* n^r since cells can at most hold one ball. For the first ball, there are n cells; for the second ball, $n-1$ cells, etc. Thus

$$\begin{aligned} N_T &= n(n-1)\cdots(n-r+1) \\ &= \frac{n!}{(n-r)!}. \end{aligned}$$

Thus

$$\begin{aligned} P &= \frac{r!}{\left(\frac{n!}{(n-r)!}\right)} \\ &= \frac{r!(n-r)!}{n!} \\ &= \binom{n}{r}^{-1}. \end{aligned}$$

40. (a) Let the tribal leaders be the cells and the rifles be the balls. Then the three tribal leaders collecting the five rifles is the analog of putting five balls into three cells.

T1	0	0	0	0	0	0	1	**1**	**1**	**1**	1	2	**2**	**2**	2	3	**3**	3	4	4	5
T2	0	1	2	3	4	5	0	**1**	**2**	**3**	4	0	**1**	**2**	3	0	**1**	2	0	1	0
T3	5	4	3	2	1	0	4	**3**	**2**	**1**	0	3	**2**	**1**	0	2	**1**	0	1	0	0

(b) These are the distributions shown in non-bold. There are fifteen such distributions.

(c) Careful here! If we count only the outcomes in bold we shall get the wrong answer i.e., 6/21=0.286. The reason this answer is wrong is that the outcomes in the columns are not equally likely. The correct answer is computed using Eq.(1.8-9) i.e.,

41. (a) The probability that a specified number appears on the face of a dice is 1/6. Hence the probability of getting three specified numbers is or 1 in 216. Hence if you win you should get \$216 for every dollar bet. But the casino payout is only \$180:1.

(b) The face value of the first dice is irrelevant. The probability that the second dice matches the first is 1/6. The probability that the third dice matches the first is 1/6. Hence the probability of getting three unspecified matches is or 1 in 36.

(c) Let E_i denote that dice $i, i = 1, 2, 3$ shows a specified number. Then the probability that (at least) two specified numbers appear is

$$\begin{aligned} & P[E_1E_2E_3^c] + P[E_1E_3E_2^c] + P[E_3E_2E_1^c] + P[E_1E_2E_3] \\ = \;& 3 \times \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \\ = \;& 0.0741, \end{aligned}$$

or about 1 in 14. So per dollar bet you should get \$14 but the casino payout is only \$10.

d.-i. The next six parts can be solved by enumeration i.e., counting. However there is a systematic procedure based on the mathematical operation of convolution that can yield all of the answers from reading a graph. The details are given in Example 3.3.-5.

d. Refer to the table below:

Dice No.1	1	2	3	4	5	6
Dice No.2	1	2	3	4	5	6
Dice No.3	1	2	3	4	5	6

We note that there are only three ways of getting a 4: 1+1+2; 1+2+1,2+1+1. Hence the

probability that the sum equals 4 is $3/(6 \times 6 \times 6) = 1/72$. Thus the fair payout should be 1:72 instead of 1:60.

e. The number of ways of getting a 5 is 6: 3+1+1; 1+3+1; 1+1+3; 2+2+1; 2+1+2; 1+2+2. Hence the probability that the sum equals 5 is $6/(6 \times 6 \times 6) = 1/36$. A fair payout would be 1:36 instead of 1:30.

f.-i. follow the same enumeration.

j. Let's think of this a series of throws. The probability that the first throw matches one of the two specified numbers is 2/6. The probability that the next throw matches a specified number is 1/6. The last throw should not match either of the numbers. Its probability is 4/6. In a throw of three dice this can happen in three ways. Hence the probability is $3 \times \frac{2}{6} \times \frac{1}{6} \times \frac{4}{6} = 1/9$ or 1:9. But the payout is only 1:5.

42. For N packets there are $N!$ ways of arranging themselves, but only one way of doing it correctly. All the packet arrangements are equally likely. Hence

$$\begin{aligned} P[\{\text{correct reassembly}\}] &= 1/N! \\ &= (3628800)^{-1} \quad \text{for } N = 10, \\ &\approx 2.76 \times 10^{-7} \end{aligned}$$

43. For three packets, there are six different arrangements $(3 \cdot 2 \cdot 1 = 6)$, but only one correct one. Hence on any try

$$\begin{aligned} P_S &\triangleq P[\{\text{success}\}] \qquad \text{and} \qquad P_F \triangleq P[\{\text{failure}\}] \\ &= \frac{1}{6} \qquad\qquad\qquad\qquad\quad = 1 - P_S = \frac{5}{6}. \end{aligned}$$

For a first correct reassembly on the n^{th} try, there must be $n-1$ failures followed by on success on the n^{th} try, thus

$$P(n) = \frac{1}{6}\left(\frac{5}{6}\right)^{n-1}, \qquad n \geq 1.$$

We note in passing that this is a valid PMF, i.e. it sums to one over its support $[1, \infty)$. To find the smallest n such that

$$\sum_{k=1}^{n} \frac{1}{6}\left(\frac{5}{6}\right)^{k-1} \geq 0.95,$$

we note that the complementary event is no successes in n trials, with probability $1 - \left(\frac{5}{6}\right)^n$. Thus we seek instead the smallest n such that $1 - \left(\frac{5}{6}\right)^n \leq 1 - 0.95 = 0.05$. Thus $n \simeq \frac{\ln(0.05)}{\ln(5/6)} \doteq 16.5$. So the answer is $n = 17$.

44.

$$\begin{aligned}
\sum_{k=0}^{n} b(k;n,p) &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-1} \\
&= (p+q)^n \\
&= (p+(1-p))^n \\
&= 1^n = 1.
\end{aligned}$$

45. (a) The probability that a BM gets destroyed is

$$\begin{aligned}
1 - P[\{\text{both AMM miss}\}] &= 1-(0.2)(0.2) \\
&= 0.96.
\end{aligned}$$

Hence for all BMs to get destroyed, we need six wins in six tries:

$$\begin{aligned}
\binom{6}{6}(0.96)^6(0.04)^0 &= (0.96)^6 \\
&\simeq 0.783.
\end{aligned}$$

(b) $P[\{\text{at least one BM gets through }\}] = 1 - P[\{\text{all are destroyed}\}] \simeq 1 - 0.783 = 0.217.$

(c)

$$\begin{aligned}
P[\{\text{exactly one gets through}\}] &= \binom{6}{5}(0.96)^5(0.04)^1 \\
&= 6(0.96)^5(0.04) \\
&\simeq 0.196.
\end{aligned}$$

46. We want to compute

$$\begin{aligned}
& P[\{\text{only one BM gets through}\}|\{\text{target destroyed}\}] \\
={}& P[\{\text{only one BM gets through}\}|\{\text{at least one BM gets through}\}] \\
={}& \frac{P[\{\text{only one BM gets through}\}, \{\text{at least one BM gets through}\}]}{P[\{\text{at least one BM gets through}\}]} \\
={}& \frac{P[\{\text{only one BM gets through}\}]}{P[\{\text{at least one BM gets through}\}]} \\
={}& \frac{0.2}{0.217} \simeq 0.922.
\end{aligned}$$

47. Let

$$\begin{aligned}
A &= \{\text{Event that a chip meets specs}\}, \\
B &= \{\text{Event that a chip needs rework}\}, \quad \text{and} \\
C &= \{\text{Event that a chip is discarded}\}.
\end{aligned}$$

We have $P[A] = 0.85, P[B] = 0.10$, and $P[C] = 0.05$. The multinomial law applies here!

(a)

$$\begin{aligned} P[\{\text{all chips meet specs}\}] &= \frac{10!}{10!0!0!}(0.85)^{10}(0.10)^0(0.05)^0 \\ &\simeq 0.197. \end{aligned}$$

(b)

$$\begin{aligned} P[\{\text{ two or more discards}\}] &= 1 - P[\{\text{no discards}\}] - P[\{\text{one discard}\}] \\ &\triangleq P_2. \end{aligned}$$

Now $P[\{\text{a chip is discarded}\}] = P[C] = 0.05$, so $P[C^c] = 0.95$, thus

$$\begin{aligned} P_2 &= 1 - \binom{10}{0}(0.95)^{10}(0.05)^0 - \binom{10}{1}(0.95)^9(0.05)^1 \\ &\simeq 1 - 0.599 - 0.315 = 0.086. \end{aligned}$$

(c)

$$\begin{aligned} P[\{\text{8 meet specs, 1 needs rework, 1 discard}\}] &= \frac{10!}{8!1!1!}(0.85)^8(0.10)^1(0.05)^1 \\ &= 0.45(0.85)^8 \\ &\simeq 0.123. \end{aligned}$$

48. Let

$$\begin{aligned} A &= \{\text{Event that call is to report fire emergency}\}, \\ B &= \{\text{Event that call is to police}\}, \quad \text{and} \\ C &= \{\text{Event that call is for ambulance}\}. \end{aligned}$$

We have $P[A] = 0.15, P[B] = 0.60, P[C] = 0.25$, and the sequence 02030202030102030202 contains 1 A, 6 Bs and 3 Cs.

(a) $P[BCBBCABCBB] = 0.15^1 \times 0.60^6 \times 0.25^3 = 1.1 \times 10^{-4}$

(b) The number of distinguishable sequences is just the multinomial coefficient

$$\frac{10!}{6!3!1!} = \frac{10 \times 9 \times 8 \times 7}{3 \times 2 \times 1} = 840$$

(c) The probability that the 10 calls involve six calls to the police, three for ambulance and one to the subdepartment:

$$\frac{10!}{6!3!1!} \times 0.15^1 \times 0.60^6 \times 0.25^3 = 0.092$$

49. We use the Poisson approximation to the binomial: Eq. 1.10-2, with $p = \frac{1}{1000} = 10^{-3}, n = 100$, and $np = 0.1$. Then

$$\begin{aligned} P[\{\text{at least one diamond is found}\}] &= 1 - P[\{\text{no diamonds are found}\}] \\ &\simeq 1 - \frac{(0.1)^0}{0!}e^{-0.1} \\ &\simeq 1 - 0.9 = 0.1. \end{aligned}$$

50. Use Eq. (1.10-7) from the text with $t = 0$, $t + \tau = 10$, and $\lambda(\tau)$ as given. This gives

$$\int_0^{10} [1 - e^{-u/10}]\, du = 10e^{-1} = 3.68.$$

Thus,

$$P[k \text{ clicks in 10 seconds }] = e^{-3.68} \frac{1}{k!}(3.68)^k.$$

51. If all the tickets are in one lottery, then $P[\text{'win'}] = \frac{50}{100} = \frac{1}{2}$. If one ticket in each of 50 lotteries, the the probability of a win in any one lottery is $p = \frac{1}{100}$, but we have 50 chances to win. Hence $P[\text{'at least one win'}] = 1 - P[\text{'no win'}]$, where

$$\begin{aligned} P[\text{'no win'}] &= \binom{50}{0} p^0 (1-p)^{50} \\ &= \binom{50}{0} \left(\frac{1}{100}\right)^0 (\frac{99}{100})^{50} \\ &\doteq 0.605. \end{aligned}$$

Hence, taking one ticket in each of 50 lotteries, $P[\text{'at least one win'}] \doteq 1 - 0.605 = 0.395 < \frac{1}{2}$.

52. If 50 tickets in one lottery, then $E[G_1] = G_1 p = 100 \cdot (\frac{1}{2}) = 50$. If one ticket in each of 50 lotteries, we would have

$$\begin{aligned} E[G_{50}] &= \sum_{i=1}^{50} 100i \binom{50}{i} \left(\frac{1}{100}\right)^i \left(\frac{99}{100}\right)^{50-i} \\ &= \frac{50}{100} 100 \left(\sum_{i'=0}^{49} \binom{49}{i'} \left(\frac{1}{100}\right)^{i'} \left(\frac{99}{100}\right)^{49-i'} \right), \quad \text{with } \ i' = i - 1, \\ &= \frac{50}{100} 100 \times 1, \quad \text{since the sum in parentheses is 1,} \\ &= 50. \end{aligned}$$

53. (a) A closed circuit can occur as

$$(A_2 A_4 \cup A_3 A_5)\, A_1 A_6 = A_1 A_2 A_4 A_6 \cup A_1 A_3 A_5 A_6.$$

(b) Now in general $P[A \cup B] = P[A] + P[B] - PAB]$, thus

$$\begin{aligned} P[\{\text{at least one closed path}\}] &= P[A_1 A_2 A_4 A_6] + P[A_1 A_3 A_5 A_6] - P[A_1 A_2 A_3 A_4 A_5 A_6] \\ &= 2p^4 - p^6 \\ &= 2p^4 (1 - \frac{1}{2} p^2). \end{aligned}$$

54. (a) Events associated with disjoint time intervals, under Poisson law, are independent. The number of cars arriving at a tollbooth in the time interval $(0, T)$ at a rate of λ per minute is such that $P[k \text{ cars arrive in } (0, T)] = e^{-\lambda T} \frac{[\lambda T]^k}{k!}$. Let us define the events:

$$\begin{aligned} A &\triangleq \{\, n_1 \text{ cars arrive in } (0, t_1)\}, \\ B &\triangleq \{\, n_2 \text{ cars arrive in } (t_1, T)\}, \text{ and} \\ C &\triangleq \{\, n_1 + n_2 \text{ cars arrive in } (0, T)\}. \end{aligned}$$

We are asked to find $P[A|C]$. From the definition of conditional probability, we know that this equals $\frac{P[AC]}{P[C]}$. The event AC is the event that n_1 cars arrive in $(0, t_1)$ and $n_1 + n_2$ cars arrive in $(0, T)$. This is the same as saying that n_1 cars arrive in $(0, t_1)$ and n_2 cars arrive in (t_1, T), which is nothing but the event AB. Therefore, $AC = AB$. But from the Poisson law (given), we know that $P[AB] = P[A]P[B]$, because A and B are events on disjoint time intervals. Therefore,

$$\begin{aligned}
P[A|C] &= \frac{P[AC]}{P[C]} = \frac{P[AB]}{P[C]} = \frac{P[A]P[B]}{P[C]} \\
&= \frac{\frac{(\lambda t_1)^{n_1} e^{-\lambda t_1}}{n_1!} \frac{(\lambda(T-t_1))^{n_2} e^{-\lambda(T-t_1)}}{n_2!}}{\frac{(\lambda T)^{n_1+n_2} e^{-\lambda T}}{(n_1+n_2)!}} \\
&= \frac{t_1^{n_1}(T-t_1)^{n_2}}{T^{n_1+n_2}} \frac{(n_1+n_2)!}{n_1! n_2!},
\end{aligned}$$

and that is independent of λ.

b. Substituting $T = 2$, $t_1 = 1$, and $n_1 = n_2 = 5$, we get

$$P[5 \text{ cars in } (0,1)|10 \text{ cars in } (0,2)] = \frac{10!}{5!5!} \frac{1}{2^{10}} \approx 0.25$$

.

55. The probability of a patient dying without the monitoring system is:

$$P_B = 0.1/2 = 0.05$$

The probability of a patient dying with the monitoring system is:

$$P(B, M) = P(B)P(M) = 0.05 \times 0.1 = 0.005$$

B and M are independent events.
Thus, Prof. X's argument is wrong.

56. (a) At each attempt, the probability of successful transmission is p^N. The repeated experiments are Bernoulli trials. Now the event $S(m)$ ={at least one successful transmission occurs in m attempts}. Also define $F(m) \triangleq$ {no successful transmission occurs in m attempts}. Then these events are mutually exclusive, so

$$\begin{aligned}
P(m) &\triangleq P[S(m)] = 1 - P[F(m)] \\
&= 1 - \binom{m}{0} (p^N)^0 (1-p^N)^m \\
&= 1 - (1-p^N)^m.
\end{aligned}$$

(b) For an individual receiver, we need the probability of at least one successful transmission in m attempts (trials). This is just the answer to part a, with p substituted for p^N, i. e. $1 - (1-p)^m$. Next consider the event $S_D(m) \triangleq$ {For every receiver, at least one successful transmission occurs in m attempts}. We have

$$P_D(m) \triangleq P[S_D(m)] = [1 - (1-p)^m]^N,$$

since there are N independent receivers.
If $p = 0.9, N = 5$, and $m = 2$, then,

$$P(2) \approx 0.832 \text{ and } P_D(2) \approx 0.951.$$

57. The sample space for the compound experiment is

$$\Omega = \{(x_1, x_2, ..., x_{100}) : 2 \leq x_i \leq 12, 1 \leq i \leq 100\}.$$

For the individual experiment with the two die, we can write the sample space as the locations (ξ_1, ξ_2) in the 6×6 table

(ξ_1, ξ_2)	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

,

where we have entered in each cell the sum of the die's upward faces. Now we set the event $A \triangleq$\{the sum is 7\} and find $P[A] = 6/36 = 1/6 \triangleq p$. As for the compound experiment consisting of $N = 100$ tries, it is seen to be Bernoulli trials with $n = 100$ and $p = 1/6$. So the answer for '10 seven's in 100 tries' is $b(7; 100, 1/6) = \begin{pmatrix} 100 \\ 10 \end{pmatrix} p^{10}(1-p)^{90}$. We can evaluate this simply using the Poisson approximation with $a = np = 100/6$. Then

$$\begin{aligned} P[10 \text{ seven's in 100 tries}] &\approx \frac{a^{10}e^{-a}}{10!} \\ &= \frac{(100/6)^{10}\, e^{-100/6}}{10!} \\ &\approx 0.0264. \end{aligned}$$

58. b) We do part (b) first. From the landlord's viewpoint, the following applies. If lease includes *free repairs*, then the cost of the two "Cloggers" versus one "NeverFail" is the same, so it doesn't matter. If repairs are not free and are the same for the "Cloggers" as for the "NeverFail," then clearly the "NeverFail" is the cheaper to lease.

a) From the tenants' point of view:

$$\begin{aligned} P[\{\text{at least one "Clogger" on}\}] &= 1 - P[\{\text{both fail}\}] \\ &= 1 - (0.4)^2 \\ &= 0.84, \text{ while} \end{aligned}$$

$$\begin{aligned} P[\{\text{"NeverFail" on}\}] &= 0.80 \\ &< 0.84. \end{aligned}$$

Therefore, the two "Cloggers" are better since at least one of them will be working 84% of the time.

59. This problem typifies the faulty reasoning exhibited by many people not familiar with probability and, in particular, the idea of *independent events*. While it is true that $P[\{2\text{ bombs on board}\}] = 10^{-4}$, the issue is $P[\{\text{terrorist bomb on board}\}|\{\text{"nervous" has a bomb on board}\}] = P[\{\text{terrorist bomb on board}\}] = 10^{-2}$, since the events $A = [\{\text{terrorist bomb on board}\}]$ and $B = [\{\text{"nervous" has a bomb on board}\}]$ are independent. Hence $P[A|B] = P[A]$; therefore, no protection!

60. Let $E = \{\text{event of successful transmission on short path}\}$;$F = \{\text{event of successful transmission on a long path}\}$. Then $P[\{\text{successful transmission}\}] = 1 - P[E^cF^c]$ and $P[E^c] = 1 - p^3$, while $P[F^c] = 1 - p^5$, where $p \triangleq 1 - q$. Therefore

$$\begin{aligned} P[\{\text{successful transmission}\}] &= 1 - P[E^cF^c] \\ &= 1 - (1-p^3)(1-p^5) \\ &= p^3 + p^5 - p^8. \end{aligned}$$

61. In this case the telephone company might find the union directive unreasonable. Here is why:

$$\begin{aligned} P[\text{overtime}] &= \sum_{n=5761}^{\infty} \frac{(720\times 8)^n}{n!} e^{-720\times 8} \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{l_1}^{l_2} e^{-\frac{1}{2}x^2} dx \\ &\quad \text{(by the Gaussian approximation to Poisson),} \end{aligned}$$

where

$$l_1 = \frac{5761 - 5760 - 0.5}{\sqrt{5760}} \quad \text{and} \quad l_2 = \frac{\infty - 5760 + 0.5}{\sqrt{5760}}.$$

Hence

$$\begin{aligned} P[\text{overtime}] &\approx \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{1}{2}x^2} dx \\ &= \operatorname{erf}(\infty) = \frac{1}{2}. \end{aligned}$$

So approximately half the time, Curtis will collect overtime.

62. The sample space for the compound experiment is

$$\Omega = \{(x_1, x_2, ..., x_{80}) : 2 \le x_i \le 4, 1 \le i \le 80\}.$$

For the individual experiment with the two die, we can write the sample space as the locations (ξ_1, ξ_2) in the 2×2 table

(ξ_1, ξ_2)	1	2
1	2	3
2	3	4

,

where we have entered in each cell the sum of the coin's upward sides. Now we set the event $A \triangleq \{\text{the sum is 2}\}$ and find $P[A] = 1/4 \triangleq p$. As for the compound experiment consisting of $N = 80$ tries, it is seen to be Bernoulli trials with $n = 80$ and $p = 1/4$. So the answer for '10 two's in 80 tries' is $b(10;80,1/4)=\begin{pmatrix}80\\10\end{pmatrix}p^{10}(1-p)^{70}$. We can evaluate an approximation simply using the Poisson approximation with $a=np=80/4=20$. Then

$$\begin{aligned}P\acute{[}10 \text{ two's in 80 tries}] &\approx \frac{a^{10}e^{-a}}{10!}\\ &= \frac{(20)^{10}\,e^{-20}}{10!}\\ &\approx 0.0058.\end{aligned}$$

Incidently the exact answer for the binomial $b(10;80,1/4)$ is 0.0028. The Poisson approximation is only marginal here since $p=1/4$ is not really $<<1$.

63. Since arrivals in disjoint intervals are independent under the Poisson law, it follows that we equivalently want the probability of 5 cars arriving in the second 2 minutes. This is given as

$$P[5 \text{ cars in 2 minutes}]=\frac{(2\lambda)^5}{5!}e^{-2\lambda}.$$

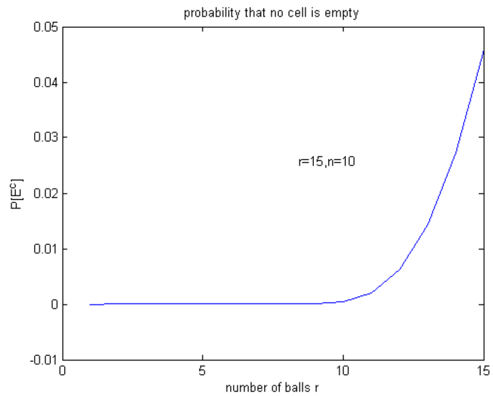
64. Unfortunately, very small. The reader should recognize that this is an occupancy problem with the candies being the "balls" and the students being the "cells." The appropriate formula is Eq. 1.8-9, which gives the probability that no cell is empty. Hence, with $r=15, n=10$,

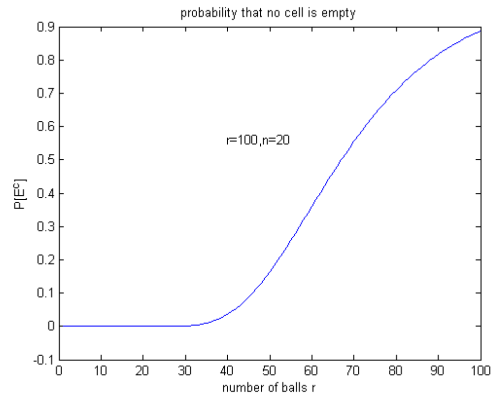
$$\begin{aligned}P[\{\text{no student is without a candy}\}] &= \sum_{i=0}^{10}\binom{10}{i}(-1)^i(1-\frac{i}{10})^{15}\\ &\approx 0.05.\end{aligned}$$

A MATLAB function to do this problem is as follows:

```
function [tries,prob] = occupancy(balls,cells)
% Here #balls=r and #cells=n. This function then
% calculates the answer to the occupancy problem in
% Example 1.8-5, specifically Eq. 1.8-9. This function
% is used in the solution to Problem 1.64 .
%
tries=1:balls; prob=zeros(1,balls);
c=zeros(1,cells); d=zeros(1,cells);
term=zeros(1,cells);
for m=1:balls
for k=1:cells
c(k)=((-1)^k)*prod(1:cells)/(prod(1:k)*prod (1:cells-k));
d(k)=(1-(k/cells))^m;
term(k)=c(k)*d(k);
end
prob (m)=1+sum(term);
end
plot(tries,prob)
xlabel('number of balls r');
ylabel('P[E^c]');
title('probability that no cell is empty');
end
```

Two example runs follow. The first is for $r = 15, n = 10$, yielding the answer to this problem at $n = 10$. The second run is for a larger case with $r = 100$ and $n = 20$.





65. These are repeated Bernoulli trials resulting in the Binomial distribution with $n = 1000$ and $p = 0.001$. Let X_i be the individual random variables, taking on value 1 for an erroneous line and 0 for an error-free line. Then we can write the sum or total of the errors as

$$Z = \sum_{i=1}^{n} X_i.$$

Then Z is Binomial with $\mu_Z = np = 1$ and $\sigma_Z^2 = npq = 0.999$. We can use the Poisson approximation to the Binomial with $a = np = 1$ here. Then

$$\begin{aligned} P[2 &\leq Z \leq 1000] = 1 - P_Z(0) - P_Z(1) \\ &\approx 1 - e^{-a} - ae^{-a} \\ &= 1 - 2e^{-1} \\ &\doteq 0.264. \end{aligned}$$

The CLT approximation gives a Normal distribution with mean $\mu = 1$ and $\sigma = \sqrt{0.999} = 0.9995$. However, it is not as accurate here since the mean μ_Z is only approximately one standard deviation away from 0, the minimum value of a Bernoulli random variable. Calculating the CLT approximate answer, we find

$$\begin{aligned} P[2 \leq Z \leq 1000] &\approx \frac{1}{\sqrt{2\pi \times 0.999}} \int_2^{1000} e^{-\frac{1}{2}\left[\frac{z-1}{\sqrt{0.999}}\right]^2} dz \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{1.0005}^{+\infty} e^{-\frac{1}{2}x^2} dx \\ &= 0.5 - \operatorname{erf}(1.0005) \\ &\doteq 0.5 - 0.341 \\ &= 0.159, \quad \text{not very accurate here.} \end{aligned}$$

66. This is a classic problem and the solution is unexpected. Let $A = \{$event that *no* two people have their birthdays on the same date$\}$ Let $B = \{$event that at least two people have their birthdays on the same date$\}$. Then $B = A^c$ and

$$\begin{aligned} P[B] &= P[A^c] \\ &= 1 - P[A], \end{aligned}$$

where

$$\begin{aligned} P[A] &= \frac{\text{no. of ways } A \text{ can occur}}{\text{no. of all possible outcomes}} \\ &= \frac{n_A}{n_T}, \end{aligned}$$

where $n_A = 365(365-1)(365-2)\cdots(365-(n-1))$ and $n_T = (365)^2$. Thus

$$\begin{aligned} P[A] &= \frac{n_A}{n_T} \\ &= 1\left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right)...\left(1-\frac{n-1}{365}\right) = \frac{1}{2}?. \end{aligned}$$

Taking logarithms, we have

$$\ln 1 + \ln\left(1-\frac{1}{365}\right) + \ln\left(1-\frac{2}{365}\right) + \cdots + \ln\left(1-\frac{n-1}{365}\right) = -0.7?.$$

Then, upon using $\ln(1-x) \simeq x$ for x small, when $n/365$ is small, we get

$$\begin{aligned} -\left(\frac{1+2+\cdots+(n-1)}{365}\right) &= -\frac{n(n-1)}{2(365)} \\ &\simeq -0.7. \end{aligned}$$

So we must set

$$\frac{n(n-1)}{2(365)} = 0.7$$

and solve for n, resulting in the quadratic equation

$$n^2 - n - 511 = 0,$$

whose positive root is $n = 22.6$. Rounding up to the next integer, we get the answer at just 23 people necessary for the probability to be one-half or greater that at least two people will have their birthday on the same date.

67. By the problem statement, we have a Binomial probability law $b(k; N, p)$ with $N = 10$ and $p = P[\text{defect}] = 0.02$. So the probability of more than one defect in the sample is given as

$$\begin{aligned} P[\text{more than 1 defect}] &= \sum_{k=2}^{10} b(k; 10, 0.02) \\ &= 1 - \sum_{k=0}^{1} b(k; 10, 0.02) \\ &= 1 - (0.98)^{10} - \binom{10}{1}(0.02)(0.98)^9 \\ &= 1 - (0.98)^{10} - 0.2(0.98)^9 \\ &= 1 - 1.18(0.98)^9 \\ &= 0.0162 \end{aligned}$$

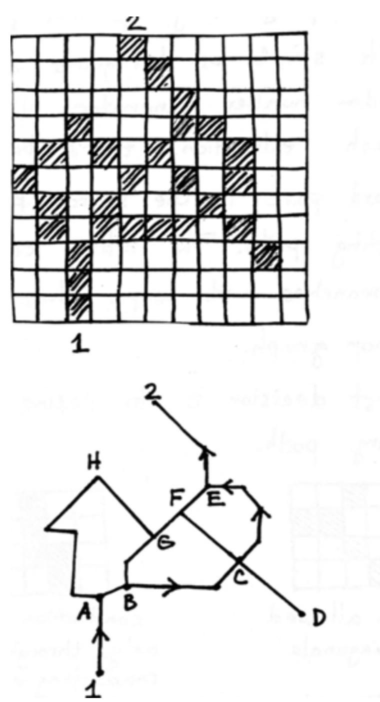
68. The programming of this problem is quite simple as it requires applying a random number generator N times for each realization ... that's basically it. The hard part is the search for a percolating path. The lattice contains both branches and loops. Thus it is netither tree, nor graph. The first decision is to define a conducting path, and there are two choices:

conduction allowed through diagonal elements

conduction allowed only through hor. and vert. elements

Two choices for conducting paths.

The time-consuming part is the search. Thus if you come to a node (junction) and the path you choose doesn't lead anywhere, you must be careful to return to the note and try the other path.



An Example.

A possible path is 1ABCE2. Note the dead-end at D and the possibility of endless looping if you are not careful. Since $N \leq 50$, relatively simple search techniques should work. A good MS thesis. Good luck!

69. Let:

$$\begin{aligned} A &= \{\text{door picked by you}\}, \\ B &= \{\text{door picked by MC}\}, \\ C &= \{\text{remaining door}\}, \text{ and} \\ D &= \{\text{door that leads to Rexis}\}. \end{aligned}$$

Then $AD = \{\text{door picked by you leads to Rexis}\}$, $BD = \{\text{door picked by MC leads to Rexis}\}$, and $CD = \{\text{remaining door leads to Rexis}\}$. Then

$$P[AD = \frac{1}{3} \quad \text{and} \quad P[BD \cup CD] = \frac{2}{3}.$$

But, since $BD \cap CD = \phi$,

$$\begin{aligned} P[BD] + P[CD] &= P[BD \cup CD] \\ &= \frac{2}{3}, \end{aligned}$$

and the MC always chooses the wrong door, so that $P[BD] = 0$, and hence $P[CD] = \frac{2}{3}$. Therefore, you should switch to door C, as it will double your probability of winning the Rexis!

70. (a) This is Bernoulli trials. Thus $P_1[E] = \binom{10}{4}\left(\frac{1}{2}\right)^{10}$ and $P_2[E] = \binom{10}{4}p^4(1-p)^6$.

(b) The likelihood ratio is given as

$$\begin{aligned} L &= P_1[E]/P_2[E] \\ &= \binom{10}{4}\left(\frac{1}{2}\right)^{10} / \binom{10}{4}p^4(1-p)^6 \\ &= \left(\frac{1}{2}\right)^{10} p^{-4}(1-p)^{-6} \\ &= \left(\frac{1}{2p}\right)^4 \left(\frac{1}{2(1-p)}\right)^6. \end{aligned}$$

Solutions to Chapter 2

1. Calculating with the binomial probability law, for the given random variable X, we obtain

$$
\begin{aligned}
P[X &= 1] = \sum_{k=0}^{2} \binom{10}{k} (.3)^k (0.7)^{10-k} \approx 0.383, \\
P[X &= 2] = \sum_{k=3}^{5} \binom{10}{k} (0.3)^k (0.7)^{10-k} \approx 0.57, \\
P[X &= 3] = \sum_{k=6}^{8} \binom{10}{k} (0.3)^k (0.7)^{10-k} \approx 0.0472, \quad \text{and} \\
P[X &= 4] = \sum_{k=9}^{10} \binom{10}{k} (0.3)^k (0.7)^{10-k} \approx 1.44 \times 10^{-4}.
\end{aligned}
$$

The CDF is plotted in Fig. 1.

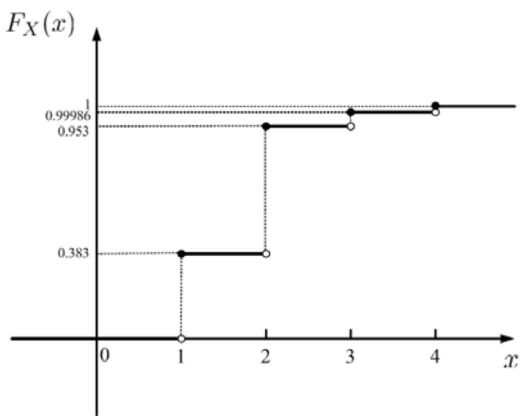


Figure 1:

2. Consider any continuous random variable. Let the outcomes be the values themselves, i.e. the random variable is an identity mapping. Then the probability of any outcome $x = \zeta$ is 0. Strictly speaking, we mean the singleton events $\{x\}$ have probability zero.

3. The cumulative distribution function (CDF) for the waiting time X is defined over $[0, \infty)$ and given as

$$
F_X(x) = \begin{cases} (x/2)^2, & 0 \le x < 1, \\ x/4, & 1 \le x < 2, \\ 1/2, & 2 \le x < 10, \\ x/20, & 10 \le x < 20, \\ 1, & 20 \le x. \end{cases}
$$

   (a) Plotting we get Fig. 2.

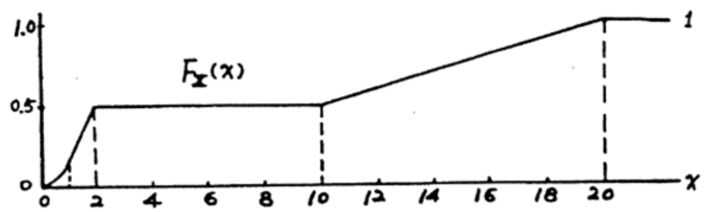


Figure 2:

(b) Taking the derivative, we get the probability density function (pdf) as

$$f_X(x) = \begin{cases} x/2, & 0 \le x < 1, \\ 1/4, & 1 \le x < 2, \\ 0, & 2 \le x < 10, \\ 1/20, & 10 \le x < 20, \\ 0, & 20 \le x. \end{cases},$$

with sketch given as:

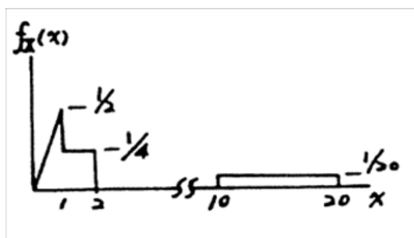


Figure 3:

We notice that

$$\begin{aligned} \int_0^\infty f_X(x)dx &= \int_0^1 \frac{x}{2}dx + \int_1^2 \frac{1}{4}dx + \int_{10}^{20} \frac{1}{20}dx \\ &= \frac{1}{4} + \frac{1}{4} + \frac{10}{20} = 1. \end{aligned}$$

(c) We compute the following probabilities:

i. [label=(0)]
ii. $P[X \ge 10] = 1 - P[X < 10] =$ [1] $1 - F_X(10) = 1/2$,
iii. $P[X < 5] = F_X(5) = 1/2$,
iv. $P[5 < X < 10] = \int_5^{10} f_X(x)dx = 0$, and
v. $P[X = 1] = 0$.

4. The point of this problem is to be careful about whether the end-points $X = b$ and $X = a$ are included in the event or not. Remember $F_X(x) \triangleq P[X \le x]$ which includes its end-point

[1] Note: $F_X(10) = P[X \le 10]$, but here the probability that $X = 10$ is zero.

at the right. Thus to compute $P[X < x]$, we must subtract from $F_X(x)$, the probability that $X = x$, i.e. $P[X = x]$.

$$\begin{aligned}
P[X < a] &= F_X(a) - P[X = a], \\
P[X \leq a] &= F_X(a), \\
P[a \leq X < b] &= F_X(b) - F_X(a) - P[X = b] + P[X = a], \\
P[a \leq X \leq b] &= F_X(b) - F_X(a) + P[X = a], \\
P[a < X \leq b] &= F_X(b) - F_X(a), \\
P[a < X < b] &= F_X(b) - F_X(a) - P[X = b].
\end{aligned}$$

5. For normalization, we integrate the probability density function (pdf) over the whole range of the random variable and equate it to 1.

   (a) *Cauchy distribution.* The pdf of a Cauchy random variable is given by

   $$f_X(x) = \frac{B}{1 + [(x - \alpha)/\beta]^2},$$

   for $\alpha < \infty$, $\beta > 0$, and $-\infty < x < \infty$. For the integration, we will substitute $\tan\theta = \frac{x-\alpha}{\beta}$, and so $d\theta = \frac{1}{\beta}\frac{1}{1+[(x-\alpha)^2/\beta]^2}$. Therefore

   $$\begin{aligned}
   \int_{-\infty}^{\infty} f_X(x)\,dx &= \int_{-\pi/2}^{\pi/2} B\beta\,d\theta \\
   &= B\beta\left(\theta|_{-\pi/2}^{\pi/2}\right) \\
   &= B\beta\pi \\
   &= 1.
   \end{aligned}$$

   Therefore, $B = \frac{1}{\beta\pi}$.

   (b) *Maxwell distribution.* The pdf of a Maxwell random variable is given by

   $$f_X(x) = \begin{cases} Bx^2 e^{-x^2/\alpha^2} & x > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{1}$$

   Integrating the pdf (substituting $y = x/\alpha$), we obtain

   $$\begin{aligned}
   \int_0^{\infty} Bx^2 e^{-x^2/\alpha^2}\,dx &= \int_0^{\infty} B\alpha^2 y^2 e^{-y^2}\alpha\,dy \\
   &= B\alpha^3\left[\frac{-1}{2} y e^{-y^2}\Big|_0^{\infty} + \frac{1}{2}\int_0^{\infty} e^{-y^2}\,dy\right] \\
   &= 0 + \frac{B\alpha^3}{2}\frac{\sqrt{\pi}}{2} \\
   &= 1.
   \end{aligned}$$

   Therefore, $B = \frac{4}{\sqrt{\pi}\alpha^3}$.

6. (a) *Beta distribution.* The pdf of the Beta distribution is given by

$$f_X(x) = \begin{cases} Bx^b(1-x)^c & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Using the formula 6.2-1 in *Handbook of Mathematical Functions* by Abramowitz and Stegun, we get

$$\begin{aligned} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx &= \int_0^\infty \frac{t^{\alpha-1}}{(1+t)^{\alpha+\beta}}\,dt \\ &= 2\int_0^{\pi/2} (\text{sin}t)^{2\alpha-1}(\text{cos}t)^{2\beta-1}\,dt. \end{aligned}$$

Now, we look at the product of the Gamma function $\Gamma(x)$ evaluated at $b+1$ and $c+1$. Substitutions used for this integration are $t = v^2, u = w^2$ and $v = r\text{sin}\ \theta, w = r\text{cos}\ \theta$.

$$\begin{aligned} &\Gamma(b+1)\Gamma(c+1) \\ &= \left[\int_0^\infty t^b e^{-t}\,dt\right]\left[\int_0^\infty u^c e^{-u}\,du\right] \\ &= \int_0^\infty \int_0^\infty t^b u^c e^{-(t+u)}\,du\,dt \\ &= 4\int_0^\infty \int_0^\infty v^{2b+1}w^{2c+1}e^{-[(v^2+w^2)]}\,dv\,dw \\ &= 4\int_0^{\pi/2} \int_0^\infty r^{2b+2c+2}(\sin\theta)^{2b+1}(\cos\theta)^{2c+1}e^{-r^2}\,rdr\,d\theta \\ &= 2\int_0^\infty e^{-r^2}(r^2)^{b+c+1}2r\,dr \int_0^{\pi/2}(\sin\theta)^{2b+1}(\cos\theta)^{2c+1}\,d\theta \\ &= \Gamma(b+c+2)2\int_0^{\pi/2}(\sin\theta)^{2b+1}(\cos\theta)^{2c+1}\,d\theta. \end{aligned}$$

Therefore, $B = \frac{\Gamma(b+c+2)}{\Gamma(b+1)\Gamma(c+1)}$.

(b) *Chi-square distribution.* The pdf of a Chi-square random variable is given by

$$f_X(x) = \begin{cases} Bx^{(n/2)-1}e^{-x/2\sigma^2} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Integrating $f_X(x)$, we get

$$\begin{aligned} \int_0^\infty f_X(x)\,dx &= \int_0^\infty Bx^{(n/2)-1}e^{-x/2\sigma^2}\,dx \\ &= \int_0^\infty B(2\sigma^2)^{(n/2)}[(x/2\sigma^2)^{(n/2)}]e^{-x/2\sigma^2}\,dx \\ &= B(2\sigma^2)^{(n/2)}\int_0^\infty [(x/2\sigma^2)^{(n/2)}]e^{-x/2\sigma^2}\,dx \\ &= B(2\sigma^2)^{(n/2)}\Gamma\left(\frac{n}{2}\right) \\ &= 1. \end{aligned}$$

Therefore, $B = \frac{1}{(2\sigma^2)^{(n/2)}\Gamma\left(\frac{n}{2}\right)}$.

7. Here we do calculations with the Normal (Gaussian) random variable of mean 0 and given variance σ^2. In notation we often indicate this as $X : N(0, \sigma^2)$. In order to calculate the probabilities $P[|X| \geq k\sigma]$ for integer values $k = 1, 2, \ldots$, we need to convert this to the standard Normal curve that is distributed as $N(0, 1)$. In particular the so-called error function is defined as

$$erf(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp(-\frac{1}{2}v^2)dv, \text{ for } x \geq 0,$$

and so only includes the right hand side of the $N(0, 1)$ distribution. Expanding $P[|X| \geq k\sigma]$ for k positive, we get $P[|X| \geq k\sigma] = P[\{X \leq -k\sigma\} \cup \{X \geq k\sigma\}]$, which is somewhat cumbersome, so instead we consider the complementary event $\{|X| < k\sigma\}$ which satisfies $P[|X| \geq k\sigma] = 1 - P[|X| < k\sigma]$. For this complementary event, we have

$$\begin{aligned} P[|X| &< k\sigma] = P[-k\sigma < x < k\sigma], \quad \text{for } k > 0, \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-k\sigma}^0 \exp(-\frac{x^2}{2\sigma^2})dx + \frac{1}{\sqrt{2\pi}} \int_0^{k\sigma} \exp(-\frac{x^2}{2\sigma^2})dx \\ &= 2\frac{1}{\sqrt{2\pi}\sigma} \int_0^{k\sigma} \exp(-\frac{x^2}{2\sigma^2})dx, \text{ by the symmetry about } v = 0. \end{aligned}$$

By making the change of variable $y = x/\sigma$, we then convert this equation into

$$\begin{aligned} P[|X| &< k\sigma] = 2\frac{1}{\sqrt{2\pi}} \int_0^k \exp(-\frac{y^2}{2})dy, \quad \text{since } dy = \frac{dx}{\sigma}, \\ &= 2\,\mathrm{erf}(k), \quad \text{for } k > 0, \end{aligned}$$

allowing us to use the standard Table 2.4-1 for $\mathrm{erf}(\cdot)$. Looking up this value and subtracting twice it from one, we get

$$\begin{aligned} k = 1, &\quad P[|X| \geq \sigma] \doteq 0.3174, \\ k = 2, &\quad P[|X| \geq 2\sigma] \doteq 0.0456, \\ k = 3, &\quad P[|X| \geq 3\sigma] \doteq 0.0026, \\ k = 4, &\quad P[|X| \geq 4\sigma] \doteq 0.0008 \approx 0. \end{aligned}$$

8. The pdf of the Rayleigh random variable is given by

$$f_X(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} u(x).$$

Note that since $f_X(x)$ is zero for negative x, $F_X(x) = 0$, for $x < 0$. Now $F_X(k\sigma) = \int_0^{k\sigma} \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}\, dx$. Substituting $y = \frac{x^2}{2\sigma^2}$ and $dy = \frac{x}{\sigma^2}$, we get

$$F_X(k\sigma) = \int_0^{k^2/2} e^{-y}\, dy = 1 - e^{-k^2/2} \qquad k = 0, 1, 2, \ldots.$$

9. For the Bernoulli random variable X, with $P_X(0) = p, P_X(1) = q$, and $q \triangleq 1 - p$, the pdf is given as

$$f_X(x) = p\delta(x) + q\delta(x - 1).$$

For the binomial random variable B with parameters n and p, we have as a function of b,

$$f_B(b) = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} \delta(b - k).$$

For the Poisson case, with mean $\mu_X = a$, we have the density

$$f_X(x) = \sum_{k=0}^{\infty} \frac{a^k}{k!} e^{-a} \delta(x-k).$$

10. This problem does some calculations with a mixed random variable. We can represent the pdf of X as

$$f_X(x) = Ae^{-x}[u(x-1) - u(x-4)] + \frac{1}{4}\delta(x-2) + \frac{1}{4}\delta(x-3).$$

(a) To find the constant A, we must integrate the pdf over all x to get 1.

$$\begin{aligned} A\int_1^4 e^{-x}dx + \frac{1}{4}\int_{-\infty}^{+\infty} \delta(x-2)dx + \frac{1}{4}\int_{-\infty}^{+\infty} \delta(x-3)dx &= 1, \\ A(e^{-1} - e^{-4}) + \frac{1}{4} + \frac{1}{4} &= 1, \end{aligned}$$

which has solution $A = \frac{1}{2}\frac{1}{e^{-1}-e^{-4}} \doteq 1.43$.

(b) Taking the running integral $\int_{-\infty}^{x} f_X(v)dv$, we get the CDF $F_X(x)$ with sketch given in Fig. 4.

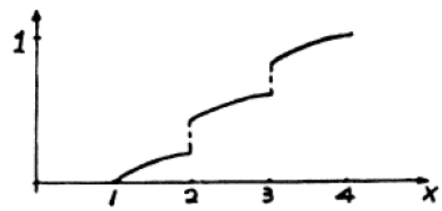


Figure 4:

Although not requested, the CDF is given analytically as

$$F_X(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{2}\frac{e^{-1}-e^{-x}}{e^{-1}-e^{-4}}, & 1 \le x < 2, \\ \frac{1}{2}\frac{e^{-1}-e^{-x}}{e^{-1}-e^{-4}} + \frac{1}{4}, & 2 \le x < 3, \\ \frac{1}{2}\frac{e^{-1}-e^{-x}}{e^{-1}-e^{-4}} + \frac{1}{2}, & 3 \le x < 4, \\ 1, & 4 \le x. \end{cases}$$

(c) We calculate the pdf as

$$f_X(x) = \frac{1}{2}\frac{e^{-x}}{e^{-1}-e^{-4}}[u(x-1) - u(x-4)] + \frac{1}{4}\delta(x-2) + \frac{1}{4}\delta(x-3).$$

So

$$\begin{aligned} P[2 \;\le\; & X < 3] = \int_2^3 \frac{1}{2}\frac{e^{-x}}{e^{-1}-e^{-4}}dx + \frac{1}{4}\int_{2^-}^{3} \delta(x-2)dx \\ &= \frac{1}{2}\frac{e^{-2}-e^{-3}}{e^{-1}-e^{-4}} + \frac{1}{4}, \end{aligned}$$

where we start the integral of the impulse at 2^- in order to pick the probability mass at $x = 2$. Note that we must include the probability mass at $x = 2$ because the event $\{2 \le X < 3\}$ includes this point.

(d) We calculate

$$\begin{aligned} P[2 \; < \; X \le 3] &= \int_2^3 \frac{1}{2}\frac{e^{-x}}{e^{-1}-e^{-4}}dx + \frac{1}{4}\int_2^{3^+} \delta(x-3)dx \\ &= \frac{1}{2}\frac{e^{-2}-e^{-3}}{e^{-1}-e^{-4}} + \frac{1}{4}, \end{aligned}$$

where we end the integral of the impulse at 3^+ to pick up the probability mass at $x = 3$.

(e) We have

$$\begin{aligned} &F_X(3) \\ &= P[X \le 3] \\ &= \int_1^3 \frac{1}{2}\frac{e^{-x}}{e^{-1}-e^{-4}}dx + \frac{1}{4}\int_1^{3^+} \delta(x-3)dx + \frac{1}{4}\int_1^{3^+} \delta(x-2)dx \\ &= \frac{1}{2}\frac{e^{-1}-e^{-3}}{e^{-1}-e^{-4}} + \frac{1}{4} + \frac{1}{4}. \end{aligned}$$

11. First we need to calculate the probability that X is less than 1 and that it is greater than 2 (area of shaded region in Fig. 5). Now

$$P[X < 1] = P[X \le 1] = F_X(1) = 1 - e^{-1}.$$

$$P[X > 2] = 1 - P[X \le 2] = 1 - F_X(2) = 1 - (1 - e^{-2}) = e^{-2}.$$

Since the events are disjoint, the probability that $X < 1$ or $X > 2$ is

$$P[\{X < 1\} \cup \{X > 2\}] = P[X < 1] + P[X > 2] = 1 - e^{-1} + e^{-2} = 0.767.$$

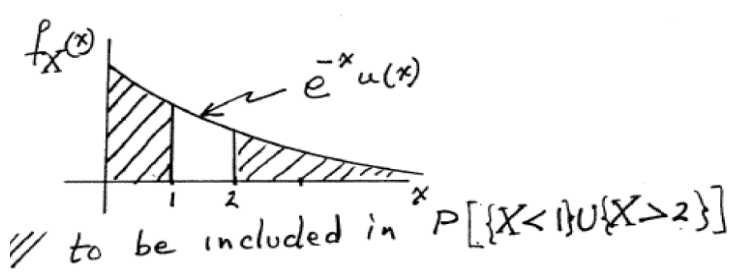


Figure 5:

12. We calculate the pdf as

$$f_X(x) = Ae^{-x}\left[u(x-1) - u(x-4)\right] + \frac{1}{4}\delta(x-2) + \frac{1}{4}\delta(x-3).$$

So

$$\begin{aligned} P[2 &< X < 4] = \int_2^4 Ae^{-x}dx + \frac{1}{4} \\ &= A(e^{-2} - e^{-4}) + \frac{1}{4}, \end{aligned}$$

where we start the integral of the impulse at 2^+ in order to not include the probability mass at $x = 2$. The overall answer then becomes $1.43(e^{-2} - e^{-4}) + 0.25$.

13. This is an example where the probability distribution is defined on the sample space which is not the elementary sample space. Normally, when we consider two coins tossed simultaneously, we consider the sample space containing two tuples of heads and tails, indicating the outcome of two tosses, i.e., we consider the sample space $\Omega = \{HH, HT, TH, TT\}$. Here we will see that we can also define probability on another set of outcomes.

    The sample space Ω contains outcomes $\zeta_1, \zeta_2, \zeta_3$ that denote outcomes of two, one, and no heads, respectively. Assuming that the coins are unbiased, we first find the probability of these outcomes.
    $P[\zeta_1] = P[\text{heads on both tosses}] = 0.5 \times 0.5 = 0.25$
    $P[\zeta_2] = P[\text{head on first, tail on second}] + P[\text{tail on first, head on second}] = 0.5 \times 0.5 + 0.5 \times 0.5 = 0.5$
    $P[\zeta_3] = P[\text{tails on both tosses}] = 0.5 \times 0.5 = 0.25$

    (a) $\{\zeta : X(\zeta) = 0, Y(\zeta) = -1\} = \zeta_2 \implies P[\{\zeta : X(\zeta) = 0, Y(\zeta) = -1\}] = P[\zeta_2] = 0.5$
    $\{\zeta : X(\zeta) = 0, Y(\zeta) = 1\} = \zeta_1 \implies P[\{\zeta : X(\zeta) = 0, Y(\zeta) = 1\}] = P[\zeta_1] = 0.25$
    $\{\zeta : X(\zeta) = 1, Y(\zeta) = -1\} = \phi \implies P[\{\zeta : X(\zeta) = 1, Y(\zeta) = -1\}] = P[\phi] = 0$
    $\{\zeta : X(\zeta) = 1, Y(\zeta) = 1\} = \zeta_3 \implies P[\{\zeta : X(\zeta) = 1, Y(\zeta) = 1\}] = P[\zeta_3] = 0.25$

    (b) Before we find the independence of X and Y, we first find the probability mass functions (pmf) of X and Y.
    $P_X[0] = P[X = 0] = P[\{\zeta_1\}] + P[\{\zeta_2\}] = 0.25 + 0.5 = 0.75$
    $P_X[1] = P[X = 1] = P[\{zeta_3\}] = 0.25.$
    $P_X[k] = 0$ for $k \neq 0, 1$.

    Similarly, $P_Y[-1] = 0.5, P_Y[1] = 0.5, P_Y[k] = 0$ for $k \neq -1, 1$.
    For independence of X, Y, we need $P_{X,Y}[a, b] = P_X[a]P_Y[b]$. For $a = 0, b = 1$, $P_{X,Y}[0, 1] = 0.25$, $P_X[0]P_Y[1] = 0.75 \times 0.5 = 0.375$.
    Hence X and Y are not independent.

14. (a) We have to integrate the given density over the full domain. We know

$$
\begin{aligned}
\int_{-\infty}^{+\infty} f_X(x)dx &= 1 \\
&= \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + A\int_{-2}^{+2} x^2 dx \\
&= \frac{1}{4} + 2A\int_{0}^{+2} x^2 dx \\
&= \frac{1}{4} + 2A\left(\frac{1}{3}x^3\Big|_0^2\right) \\
&= \frac{1}{4} + 2A\frac{8}{3}.
\end{aligned}
$$

Hence $A = 9/64$.

(b) A labeled plot appears below in Fig. 6. Note the jumps occurring at the impulse locations in the density. Also note the slope of the distribution function is given by the density function in the smooth regions.

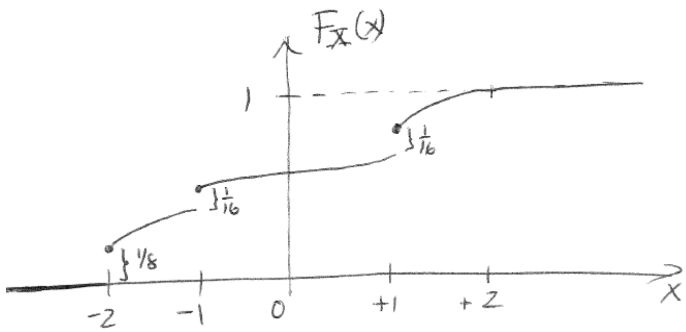


Figure 6:

(c) We proceed as follows

$$
\begin{aligned}
F_X(1) &= \int_{-\infty}^{1} f_X(x)dx \\
&= \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \frac{9}{64}\int_{-2}^{+1} x^2 dx \\
&= \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \left(\frac{3}{8} + \frac{9}{64}\int_{0}^{+1} x^2 dx\right) \\
&\qquad \text{(using the symmetry of } x^2\text{)} \\
&= \frac{1}{4} + \left(\frac{3}{8} + \frac{9}{64}\left(\frac{1}{3}x^3\Big|_0^1\right)\right) \\
&= \frac{1}{4} + \frac{3}{8} + \frac{9}{64}\frac{1}{3} = \frac{43}{64}.
\end{aligned}
$$

(d) We can calculate

$$\begin{aligned} P[-1 &< X \leq 2] = \int_{-\infty}^{2} f_X(x)dx \\ &= \frac{1}{16} + \frac{9}{64}\int_{-1}^{+2} x^2 dx. \end{aligned}$$

But the easier way is to realize that $\int_{-1}^{+2} x^2 dx = \int_{-2}^{+1} x^2 dx$ (because of symmetry) which was needed in part (c). Then, by just counting the one relevant impulse area, we can write

$$P[-1 < X \leq 2] = \frac{1}{16} + \frac{3}{8} + \frac{3}{64} = \frac{31}{64}.$$

15. First we calculate the probability that X is even. Now it is binomial distributed with parameters $n = 4$ and $p = 0.5$, i.e. $b(k; 4, 0.5), 0 \leq k \leq 4$, thus

$$\begin{aligned} P[\{X = \text{even}\}] &= b(0;4,0.5) + b(2;4,0.5) + b(4;4,0.5) \\ &= \binom{4}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4 + \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 + \binom{4}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0 \\ &= 1 \times \left(\frac{1}{2}\right)^4 + 6 \times \left(\frac{1}{2}\right)^4 + 1 \times \left(\frac{1}{2}\right)^4 = \frac{1}{2}. \end{aligned}$$

Now, the conditional probability is given as

$$\begin{aligned} P[\{X &= k\}|\{X = \text{even}\}] = \frac{P[\{X = k\} \cap \{X = \text{even}\}]}{P[\{X = \text{even}\}]} \\ &= \begin{cases} 2\frac{1}{16} = \frac{1}{8}, & k = 0, \\ 0, & k = 1, \\ 2\frac{6}{16} = \frac{6}{8}, & k = 2, \\ 0, & k = 3, \\ 2\frac{1}{16} = \frac{1}{8}, & k = 4, \end{cases} \end{aligned}$$

where we have used the fact that the joint event

$$\{X = k\} \cap \{X = \text{even}\} = \begin{cases} \{X = k\}, & k \text{ even}, \\ \phi, & k \text{ odd}. \end{cases}$$

16. The marginal distribution function of random variable N is given by

$$F_N(n) = F_{W,N}(+\infty, n) = \begin{cases} 0, & n < 0, \\ \frac{n}{10}, & 0 \leq n < 5, \\ \frac{n}{10}, & 5 \leq n < 10, \\ 1, & n \geq 10. \end{cases}$$

The pmf is given by

$$P_N(n) = F_N(n) - F_N(n-1) = \begin{cases} 0, & n \leq 0, \\ \frac{1}{10}, & 0 < n \leq 10, \\ 0, & n > 10. \end{cases}$$

The conditional probability density function:

$$\begin{aligned} P[W \le w, N = n] &= F_{W,N}(w,n) - F_{W,N}(w,n-1) \\ &= u(w) \begin{cases} 0, & n \le 0 \\ (1 - e^{-w/\mu_0})\frac{1}{10}, & 1 \le n \le 5 \\ (1 - e^{-w/\mu_1})\frac{1}{10}, & 5 < n \le 10 \\ 0, & n > 10 \end{cases}. \end{aligned}$$

Note that $F_W(w|N = n) = P[W \le w|N = n]$ is not defined for $n \le 0$ or $n > 10$, because for these n, $P[N = n] = 0$. Therefore,

$$\begin{aligned} F_W(w|N = n) &= \frac{P[W \le w, N = n]}{P_N(n)} \\ &= u(w) \begin{cases} (1 - e^{-w/\mu_0}), & 1 \le n \le 5 \\ (1 - e^{-w/\mu_1}), & 5 < n \le 10 \end{cases}. \end{aligned}$$

Hence,

$$f_W(w|N = n) = u(w) \begin{cases} \frac{1}{\mu_0} e^{-w/\mu_0}, & 1 \le n \le 5 \\ \frac{1}{\mu_1} e^{-w/\mu_1}, & 5 < n \le 10. \end{cases}$$

17. Let the number of bulbs produced by A and B be n_A and n_B respectively. We have $n_A + n_B = n$, and n is the total number of the bulbs. So $P[A] = \frac{n_A}{n} = \frac{1}{4}$ and $P[B] = \frac{n_B}{n} = \frac{3}{4}$. Since we have

$$F_X(x|A) = (1 - e^{-0.2x})u(x), \quad F_X(x|B) = (1 - e^{-0.5x})u(x),$$

then

$$\begin{aligned} F_X(x) &= F_X(x|A)P(A) + F_X(x|B)P(B) \\ &= \frac{1}{4}(1 - e^{-0.2x})u(x) + \frac{3}{4}(1 - e^{-0.5x})u(x). \end{aligned}$$

So

$$F(2) = \frac{1}{4}(1 - e^{-0.2\times 2}) + \frac{3}{4}(1 - e^{-0.5\times 2}) = 0.56,$$

$$F(5) = \frac{1}{4}(1 - e^{-0.2\times 5}) + \frac{3}{4}(1 - e^{-0.5\times 5}) = 0.85,$$

$$F(7) = \frac{1}{4}(1 - e^{-0.2\times 7}) + \frac{3}{4}(1 - e^{-0.5\times 7}) = 0.92.$$

Then $P[\text{burns at least 2 months}] = 1 - F(2) = 0.44$, $P[\text{burns at least 5 months}] = 1 - F(5) = 0.15$ and $P[\text{burns at least 7 months}] = 1 - F(7) = 0.08$.

18. Given the event $A \triangleq \{b < X \le a\}$, for $b < a$, we calculate $F_X(x|A)$.

    i) $x \le b : F_X(x|A) = 0$, since the joint event $\{X \le b\} \cap A = \phi$.

    ii) $x > a : F_X(x|A) = 1$, since the joint event $\{X \le a\} \cap A = A$, so the conditional probability of $\{X \le a\}$, given A, is one.

iii) $b < x \leq a$: Here we must calculate the actual intersection of the two sets $\{X \leq a\}$ and $A = \{b < X \leq a\}$. Since $b < x \leq a$, we get $\{X \leq a\} \cap A = \{X \leq x\} \cap \{b < X \leq a\} = \{b < X < x\}$. We can then calculate the conditional probability

$$\begin{aligned} F_X(x|A) &= \frac{P[\{X \leq x\} \cap A]}{P[A]} \\ &= \frac{P[\{b < X < x\}]}{P[A]} \\ &= \frac{F_X(x) - F_X(b)}{F_X(a) - F_X(b)}, \quad \text{for } b < x \leq a. \end{aligned}$$

19. In order to get $P[Y = k]$, we can consider $P[Y = k|X = x]$ first and then do the integral over all x.

$$\begin{aligned} P[Y = k] &= \int_{-\infty}^{\infty} P[Y = k|X = x] f_X(x) dx \\ &= \frac{1}{5} \int_0^5 \frac{x^k e^{-x}}{k!} dx = \frac{1}{5k!} \int_0^5 x^k e^{-x} dx \end{aligned}.$$

for $k = 0$:

$$P[Y = 0] = \frac{1}{5}(1 - \frac{1}{0!} e^{-5})$$

for $k = 1$:

$$P[Y = 1] = \frac{1}{5}(1 - \frac{1}{0!} e^{-5} - \frac{5^1}{1!} e^{-5})$$

for $k = 2$:

$$P[Y = 2] = \frac{1}{5}\left(1 - \frac{1}{0!} e^{-5} - \frac{5^1}{1!} e^{-5} - \frac{5^2}{2!} e^{-5}\right)$$

for general k:

$$P[Y = k] = \frac{1}{5}\left(1 - \frac{1}{0!} e^{-5} - \frac{5^1}{1!} e^{-5} \ldots - \frac{5^k}{k!} e^{-5}\right), \qquad k \geq 0.$$

20. (a) The pmf of X is binomial with $n = 8$ and $p = q = 0.5$, i.e.

$$P_X(k) \triangleq P[X = k] = b(k; 8, 0.5).$$

This is because the 8 votes are independent, each with $p = 0.5$ chance of being favorable. They are thus Bernoulli trials, which leads to the binomial distribution in the binary case. We note that since $p = 0.5$, the distribution will be symmetric about $X = k = 4$.

(b) We must find the conditional PDF $F_X(x|A)$ for the range $-1 \leq x \leq 10$. Now

$$\begin{aligned} F_X(x|A) &\triangleq P[X \leq x|A] = \frac{P[\{X \leq x\}\} \cap \{X > 4\}]}{P[X > 4]} \\ &= \frac{P[4 < X \leq x]}{P[X > 4]}. \end{aligned}$$

Since X is binomially distributed, we have

$$\begin{aligned} P[X > 4] &= \left(\frac{1}{2}\right)^8 \sum_{k=5}^{8} \binom{8}{k} \\ &= \frac{1}{2}\left(1 - \binom{8}{4}\left(\frac{1}{2}\right)^8\right) \\ &= \frac{93}{256}, \end{aligned}$$

where the second to last line is by symmetry of this binomial distribution about $k = 4$. Turning to the numerator, we have

$$P[4 < X \leq x] = \begin{cases} 0, & x < 5 \\ \left(\frac{1}{2}\right)^8 \sum_{k=5}^{8} \binom{8}{k} u(x-k), & x \geq 5. \end{cases}$$

Then

$$\begin{aligned} F_X(x|A) &= \begin{cases} 0, & x < 5 \\ \frac{256}{93}\left(\frac{1}{2}\right)^8 \sum_{k=5}^{8} \binom{8}{k} u(x-k), & x \geq 5. \end{cases} \\ &= \left(\frac{1}{93}\sum_{k=5}^{8}\binom{8}{k} u(x-k)\right) u(x-5) \\ &= \left(\frac{1}{93}\sum_{k=5}^{\lfloor x \rfloor}\binom{8}{k}\right) u(x-5), \end{aligned}$$

where $\lfloor x \rfloor$ denotes the least integer function

$$\lfloor x \rfloor \triangleq x \text{ rounded } \textit{down} \text{ to next integer.}$$

Calculating, we determine

$$\frac{1}{93}\binom{8}{k} = \begin{cases} \frac{56}{93}, & k = 5, \\ \frac{28}{93}, & k = 6, \\ \frac{8}{93}, & k = 7, \\ \frac{1}{93}, & k = 8, \end{cases}$$

and thus

$$F_X(x|A) = \begin{cases} \frac{56}{93}, & k = 5, \\ \frac{84}{93}, & k = 6, \\ \frac{92}{93}, & k = 7, \\ 1, & k = 8. \end{cases}$$

Now for $x < 5$, $F_X(x|A) = 0$, and for $x > 8$, $F_X(x|A) = 1$, so we have the plot of Figure 1.

(c) From the calculations done above and from the definition

$$\begin{aligned} f_X(x|A) &= \frac{dF_X(x|A)}{dx} \\ &= \frac{1}{93}\sum_{k=5}^{8}\binom{8}{k}\delta(x-k) \\ &= \frac{56}{93}\delta(x-5) + \frac{28}{93}\delta(x-6) + \frac{8}{93}\delta(x-7) + \frac{1}{93}\delta(x-8), \end{aligned}$$

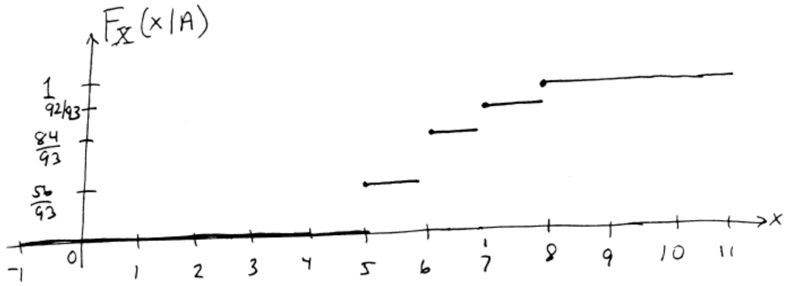


Figure 7:

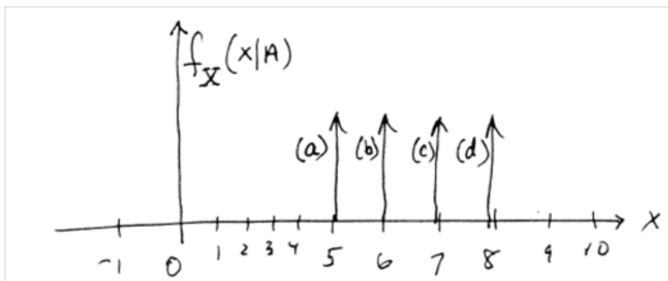


Figure 8:

with plot of Figure 2. Note we write the areas of the impulses in parentheses. In this figure, $a = 56/93$, $b = 28/93$, $c = 8/93$, and $d = 1/93$.

(d) Using Bayes' rule, we have

$$\begin{aligned} P[4 \leq X \leq 5|A] &= \frac{P[4 < X \leq 5]}{P[X > 4]} \\ &= \frac{P[X = 5]}{P[X > 4]} \\ &= \frac{\left(\frac{1}{2}\right)^8 \binom{8}{5}}{\frac{93}{2^8}} = \frac{1}{93}\binom{8}{5} = \frac{56}{93}. \end{aligned}$$

21. The random variables X and Y have joint probability density function (pdf)

$$f_{X,Y}(x,y) = \begin{cases} \frac{3}{4}x^2(1-y), & 0 \leq x \leq 2, 0 \leq y \leq 1, \\ 0, & \text{else.} \end{cases}$$

(a) To find $P[X \leq 0.5]$, we start with

$$\begin{aligned}
P[X &\leq 0.5] = \int_{-\infty}^{0.5}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dxdy \\
&= \int_0^{0.5}\int_0^1 \frac{3}{4}x^2(1-y)dxdy \\
&= \frac{3}{4}\left(\int_0^{0.5} x^2dx\right)\left(\int_0^1 (1-y)dy\right) \\
&= \frac{3}{4}\left(\frac{x^3}{3}|_0^{0.5}\right)\left((y-\frac{y^2}{2})|_0^1\right) \\
&= \frac{3}{4}\frac{1}{24}(1-\frac{1}{2}) = \frac{1}{64}.
\end{aligned}$$

.

(b) By definition

$$\begin{aligned}
F_Y(0.5) &= P[Y \leq 0.5] \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{0.5} f_{X,Y}(x,y)dxdy \\
&= \int_0^2\int_0^{0.5} \frac{3}{4}x^2(1-y)dxdy \\
&= \frac{3}{4}\left(\frac{x^3}{3}|_0^2\right)\left((y-\frac{y^2}{2})|_0^{0.5}\right) \\
&= \frac{3}{4}\frac{8}{3}\left(\frac{1}{2}-\frac{1}{8}\right) = \frac{3}{4}.
\end{aligned}$$

(c) To find $P[X \leq 0.5|Y \leq 0.5]$, we note that X and Y are independent random variables, so the answer is the same as in part a), namely $P[X \leq 0.5|Y \leq 0.5] = P[X \leq 0.5] = \frac{1}{64}$. However, we can also calculate directly,

$$\begin{aligned}
P[X &\leq 0.5|Y \leq 0.5] = \frac{P[X \leq 0.5, Y \leq 0.5]}{P[Y \leq 0.5]} \\
&= \int_0^{0.5}\int_0^{0.5} \frac{3}{4}x^2(1-y)dxdy/\left(\frac{3}{4}\right) \\
&= \frac{3}{4}\frac{1}{24}\left(\frac{1}{2}-\frac{1}{8}\right)/\left(\frac{3}{4}\right) = \frac{1}{64}.
\end{aligned}$$

(d) Here, we can note again that X and Y are independent random variables for the given joint pdf, and thus

$$\begin{aligned}
P[Y &\leq 0.5|X \leq 0.5] = P[Y \leq 0.5] \\
&= \frac{3}{4} \quad \text{from part b).}
\end{aligned}$$

22. To check for independence, we need to look at the marginal pdfs of X and Y. How do we find the pdf's? We can use the property that the pdf must integrate to 1. Say $f_X(x) = Ae^{-\frac{1}{2}(\frac{x}{3})^2}u(x)$, and $\int_0^\infty f_X(x)dx = 1$, we find $A = \frac{2}{3\sqrt{2\pi}}$. Similarly, $f_Y(y) = Be^{-\frac{1}{2}(\frac{x}{2})^2}u(y)$,

and $\int_0^\infty f_Y(y)dy = 1$, so $B = \frac{2}{2\sqrt{2\pi}}$. Multiplying the two marginal pdfs, we see that the product is indeed equal to joint pdf; i.e., $f_X(x)f_Y(y) = f_{X,Y}(x,y)$. Therefore, X and Y are independent random variables; their joint probability factors and hence $P[0 < X \le 3, 0 < Y \le 2] = P[0 < X \le 3]P[0 < Y \le 2]$. Thus

$$
\begin{aligned}
P[0 < X \le 3] &= \int_{-3}^{3} \frac{2}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{3}\right)^2} dx \\
&= 2 \times \frac{2}{3\sqrt{2\pi}} \int_0^3 e^{-\frac{1}{2}\left(\frac{x}{3}\right)^2} dx = 2erf(1),
\end{aligned}
$$

$$
\begin{aligned}
P[0 < Y \le 2] &= \int_{-2}^{2} \frac{2}{2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y}{2}\right)^2} dy \\
&= 2 \times \frac{2}{2\sqrt{2\pi}} \int_0^2 e^{-\frac{1}{2}\left(\frac{y}{2}\right)^2} dy = 2erf(1).
\end{aligned}
$$

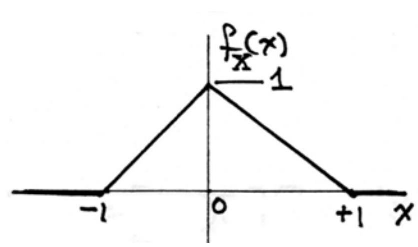
So

$$
\begin{gathered}
P[0 < X \le 3, 0 < Y \le 2] = P[0 < X \le 3]P[0 < Y \le 2] \\
= 2\,\mathrm{erf}(1) \times 2\,\mathrm{erf}(1) = 4\,\mathrm{erf}(1)^2 \doteq 0.466.
\end{gathered}
$$

23. (a) Since

$$
\begin{aligned}
1 &= \int_{-\infty}^{+\infty} f_X(x)dx \\
&= A\int_{-1}^{0}(1+x)dx + A\int_0^{+1}(1-x)dx \\
&= A\left(x + \frac{x^2}{2}\right)\Big|_{-1}^{0} + A\left(x - \frac{x^2}{2}\right)\Big|_0^1 \\
&= A\left(\frac{1}{2} + \frac{1}{2}\right) \\
&= A.
\end{aligned}
$$

Thus $A = 1$ and f_X is plotted as



(b) $F_X(x) = 0$ for $x \leq -1$. Then for $-1 < x \leq 0$, we calculate

$$\begin{aligned} F_X(x) &= \int_{-1}^{x} (1+v)dv \\ &= \left(v + \frac{v^2}{2}\right)\Big|_{-1}^{x} \\ &= x + \frac{x^2}{2} - \left(-1 + \frac{(-1)^2}{2}\right) \\ &= x + \frac{x^2}{2} + \frac{1}{2}. \end{aligned}$$

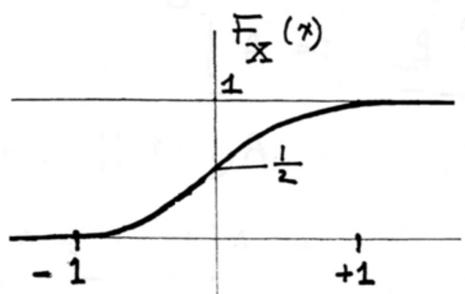
We note that $F_X(x) = \frac{1}{2}$. Then for $0 < x \leq 1$, we calculate

$$\begin{aligned} F_X(x) &= \frac{1}{2} + \int_{0}^{x} (1-v)dv \\ &= \frac{1}{2} + \left(v - \frac{v^2}{2}\right)\Big|_{0}^{x} \\ &= \frac{1}{2} + x - \frac{x^2}{2}. \end{aligned}$$

Note that $F_X(x) = 1$ for $x \geq 1$ since $\int_{-1}^{+1} f_X(x)dx = 1$. Putting all the results together, we get

$$F_X(x) = \begin{cases} 0, & x \leq -1, \\ \frac{1}{2} + x + \frac{x^2}{2}, & -1 < x \leq 0, \\ \frac{1}{2} + x - \frac{x^2}{2}, & 0 < x \leq 1 \\ 1, & x > 1. \end{cases}$$

The sketch of F_X is shown below.



(c) $P[X > b] = 1 - F_X(b) = \frac{1}{2}F_X(b)$, which gives $F_X(b) = \frac{2}{3}$, therefore $b \in (0,1)$. In this interval $F_X(b) = \frac{1}{2} + b - \frac{b^2}{2}$, so we have the quadratic equation

$$3b^2 - 6b + 1 = 0,$$

which is solved by roots $b_{1,2} = 1 \pm \sqrt{\frac{2}{3}}$. The root in $(0,1)$ is then $b = 1 - \sqrt{\frac{2}{3}} \simeq 0.185$.

24. The general expression is given as:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2\sigma^2(1-\rho^2)}(x^2 + y^2 - 2\rho xy)\right).$$

If $\rho = 0$ and $\sigma = 1$, then this becomes

$$\begin{aligned} f_{XY}(x,y) &= \frac{1}{2\pi}\exp\left(\frac{-1}{2}(x^2+y^2)\right) \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2} \\ &= f_X(x)f_Y(y). \end{aligned}$$

The desired joint probability can be calculated as

$$\begin{aligned} P\left[-\frac{1}{2} < X \le \frac{1}{2}, -\frac{1}{2} < Y \le \frac{1}{2}\right] &= P\left[-\frac{1}{2} < X \le \frac{1}{2}\right] P\left[-\frac{1}{2} < Y \le \frac{1}{2}\right] \\ &= 2\,\mathrm{erf}\left(\frac{1}{2}\right) 2\,\mathrm{erf}\left(\frac{1}{2}\right) \\ &= \left[2\,\mathrm{erf}\left(\frac{1}{2}\right)\right]^2 \\ &\simeq 0.144. \end{aligned}$$

25. We use Bayes' formula for pdf's:

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}.$$

We have

$$f_X(x) = \frac{1}{2}\mathrm{rect}(\frac{x}{2}).$$

Then

$$\begin{aligned} f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx &= \int_{-\infty}^{\infty} f_{Y|X}(y|x)f_X(x)dx \\ &= \int_{-1}^{1}\frac{1}{2}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y-x)^2}{2\sigma^2}\right]dx. \end{aligned}$$

Let $\xi = \frac{x-y}{\sigma}$, then $d\xi = \frac{dx}{\sigma}$ and we obtain

$$f_Y(y) = \frac{1}{2}\int_{\frac{-1-y}{\sigma}}^{\frac{1-y}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\xi^2}d\xi = \frac{1}{2}\left[\mathrm{erf}(\frac{1-y}{\sigma}) - \mathrm{erf}(\frac{-1-y}{\sigma})\right].$$

But $erf(x) = -erf(-x)$, hence

$$f_Y(y) = \frac{1}{2}\left[\mathrm{erf}(\frac{1+y}{\sigma}) - \mathrm{erf}(\frac{y-1}{\sigma})\right].$$

Then finally

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)} = \frac{\frac{1}{\sqrt{2\pi}\ \sigma}\exp[-\frac{(y-x)^2}{2\sigma^2}]\mathrm{rect}(\frac{x}{2})}{\mathrm{erf}(\frac{1+y}{\sigma}) - \mathrm{erf}(\frac{y-1}{\sigma})}.$$

26. We start off with the general relation for conditional probability densities

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)} \\
&= \frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(y-x)^2/2\sigma^2}\left(\frac{1}{2}\delta(x-1)+\frac{1}{2}\delta(x+1)\right)}{f_Y(y)}.
\end{aligned}$$

Next, we find the denominator as

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{+\infty} f_{Y,X}(y,x)dx \\
&= \int_{-\infty}^{+\infty} f_{Y|X}(y|x)f_X(x)dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(y-x)^2/2\sigma^2}\left(\frac{1}{2}\delta(x-1)+\frac{1}{2}\delta(x+1)\right)dx,
\end{aligned}$$

and so, combining this result with the one above, we have

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(y-x)^2/2\sigma^2}\left(\frac{1}{2}\delta(x-1)+\frac{1}{2}\delta(x+1)\right)}{f_Y(y)} \\
&= \frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(y-x)^2/2\sigma^2}\left(\frac{1}{2}\delta(x-1)+\frac{1}{2}\delta(x+1)\right)}{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(y-x)^2/2\sigma^2}\left(\frac{1}{2}\delta(x-1)+\frac{1}{2}\delta(x+1)\right)dx} \\
&= \frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(y-x)^2/2\sigma^2}\left(\frac{1}{2}\delta(x-1)+\frac{1}{2}\delta(x+1)\right)}{\frac{1}{\sqrt{2\pi\sigma^2}}\left(\frac{1}{2}e^{-(y-1)^2/2\sigma^2}+\frac{1}{2}e^{-(y+1)^2/2\sigma^2}\right)} \\
&= \frac{e^{-(y-x)^2/2\sigma^2}\left(\frac{1}{2}\delta(x-1)+\frac{1}{2}\delta(x+1)\right)}{\frac{1}{2}e^{-(y-1)^2/2\sigma^2}+\frac{1}{2}e^{-(y+1)^2/2\sigma^2}}, \quad \text{or equivalently} \\
&= \frac{e^{-(y-x)^2/2\sigma^2}}{e^{-(y-1)^2/2\sigma^2}+e^{-(y+1)^2/2\sigma^2}}\left(\delta(x-1)+\delta(x+1)\right).
\end{aligned}$$

Note, we could have eliminated a few steps in our solution by starting from the Total Probability Theorem for density functions, from which we can write directly

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{+\infty} f_{Y|X}(y|x)f_X(x)dx}.$$

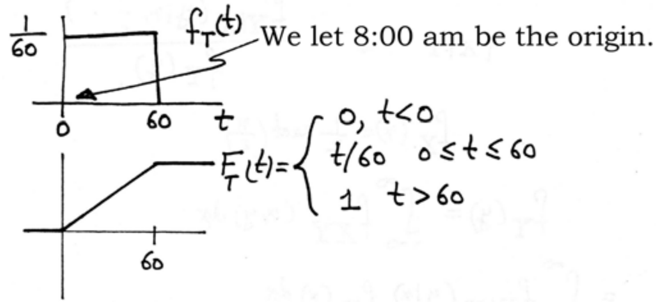
If the question had asked instead for the conditional probability mass function (PMF) $P_{X|Y}$, the answer would have been

$$P_{X|Y}(x|y) = \begin{cases} \frac{e^{-(y-x)^2/2\sigma^2}}{e^{-(y-1)^2/2\sigma^2}+e^{-(y+1)^2/2\sigma^2}}, & x=\pm 1, \\ 0, & \text{else} \end{cases},$$

as can be easily obtained by integrating the conditional density found above.

27.

Let T be the prof's arrival time.



$$
\begin{aligned}
P[A] &= P[T > 30] = 1 - F_T(30) \\
P[B] &= P[T \le 31] = F_T(31) \\
P[AB] &= P[30 < T \le 31] = F_T(31) - F_T(30) \\
P[B|A] = \frac{P[AB]}{P[A]} &= \frac{F_T(31) - F_T(30)}{1 - F_T(30)} \\
&= \frac{\frac{31-30}{60}}{\frac{60-30}{60}} = \frac{1}{30}.
\end{aligned}
$$

$$
\begin{aligned}
P[A|B] &= \frac{P[AB]}{P[B]} = \frac{F_T(31) - F_T(30)}{F_T(31)} \\
&= \frac{\frac{31-30}{60}}{\frac{31}{60}} = \frac{1}{31}.
\end{aligned}
$$

28. (a)

$$
\begin{aligned}
1 &= \int_{-\infty}^{+\infty} f_X(x)dx \\
&= c\int_0^{\infty} e^{-2x}dx \\
&= c\left(\frac{e^{-2x}}{-2}\bigg|_0^{\infty} \right) \\
&= c\left(0 - -\frac{1}{2}\right) \\
&= c/2,
\end{aligned}
$$

thus we must have $c = 2$.

(b) For $x > 0, a > 0$, we can write

$$\begin{aligned}
P[X \geq x+a] &= 2\int_{x+a}^{\infty} e^{-2v} dv \\
&= 2\left(\frac{e^{-2v}}{-2}\bigg|_{x+a}^{\infty}\right) \\
&= 2\left(0 - -\frac{e^{-2(x+a)}}{2}\right) \\
&= e^{-2(x+a)}, \quad \text{with } x > 0, a > 0.
\end{aligned}$$

(c)

$$P[X \geq x+a | X > a] = \frac{P[X \geq x+a, X > a]}{P[X > a]},$$

but note that for $x > 0$, the event $\{X \geq x+a\}$ is a subset of the event $\{X > a\}$, so $\{X \geq x+a\} \cap \{X > a\} = \{X \geq x+a\}$, and hence $P[X \geq x+a, X > a] = P[X \geq x+a]$, thus we have, for $x > 0$,

$$\begin{aligned}
P[X \geq x+a | X > a] &= \frac{P[X \geq x+a, X > a]}{P[X > a]} \\
&= \frac{P[X \geq x+a]}{P[X > a]} \\
&= \frac{e^{-2(x+a)}}{e^{-2a}} \\
&= e^{-2x}, \quad \text{independent of } a\ !
\end{aligned}$$

Since this conditional probability is (functionally) independent of the variable a, the *memory* of a had been lost.

29. We need to solve for y in

$$1 - e^{-y} \geq 0.95.$$

But this implies that

$$\begin{aligned}
y &\geq -\ln(0.005) \\
&\doteq 2.996.
\end{aligned}$$

Thus, $y = 3$ should do.

30. This is a rather classic problem in detection theory.

$$\begin{aligned}
P[A|M] &= P[X \geq 0.5|M] \\
&= \frac{1}{\sqrt{2\pi}}\int_{0.5}^{\infty} e^{-\frac{1}{2}(x-1)^2} dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-0.5}^{\infty} e^{-\frac{1}{2}y^2} dy, \quad \text{with } y \triangleq x-1, \\
&= \frac{1}{2} + \operatorname{erf}(0.5) \doteq 0.69.
\end{aligned}$$

Then

$$\begin{aligned} P[A|M^c] &= P[X \geq 0.5|M^c] \\ &= \frac{1}{\sqrt{2\pi}} \int_{0.5}^{\infty} e^{-\frac{1}{2}x^2} dx \\ &= \frac{1}{2} - \mathrm{erf}(0.5) \doteq 0.31, \end{aligned}$$

$$\begin{aligned} P[A^c|M^c] &= P[X < 0.5|M^c] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0.5} e^{-\frac{1}{2}x^2} dx \\ &= \frac{1}{2} + \mathrm{erf}(0.5) \doteq 0.69, \end{aligned}$$

and

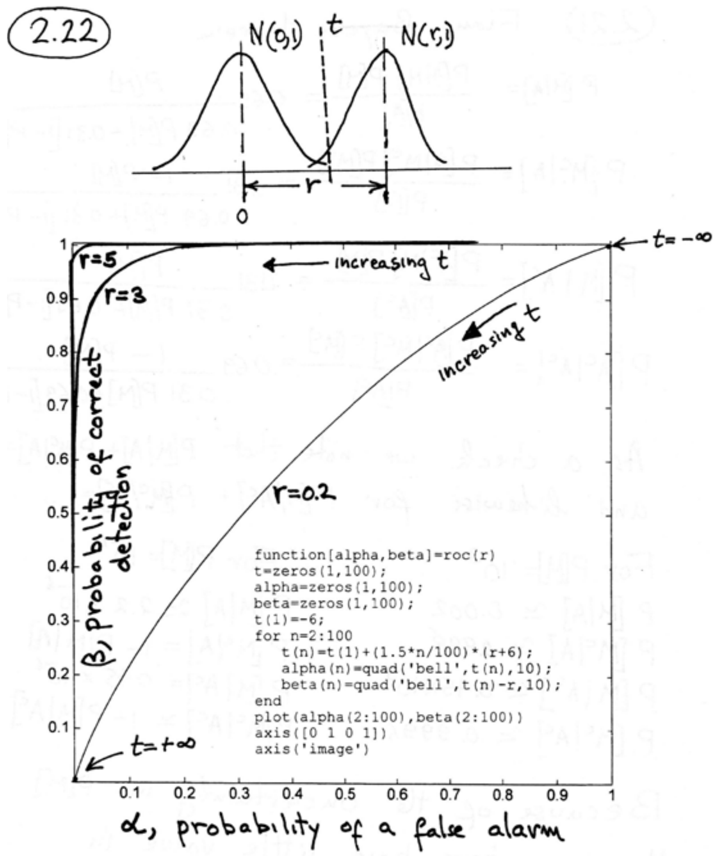
$$\begin{aligned} P[A^c|M] &= P[X < 0.5|M] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0.5} e^{-\frac{1}{2}(x-1)^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-0.5} e^{-\frac{1}{2}y^2} dy, \qquad \text{again with } y \triangleq x-1, \\ &= \frac{1}{2} - \mathrm{erf}(0.5) \doteq 0.31. \end{aligned}$$

31. From Bayes' Theorem

$$\begin{aligned} P[M|A] &= \frac{P[A|M]P[M]}{P[A]} = 0.69 \frac{P[M]}{0.69P[M] + 0.31(1-P[M])}, \\ P[M^c|A] &= \frac{P[A|M^c]P[M^c]}{P[A]} = 0.31 \frac{P[M^c]}{0.69P[M^c] + 0.31(1-P[M^c])}, \\ P[M|A^c] &= \frac{P[A^c|M]P[M]}{P[A^c]} = 0.31 \frac{P[M]}{0.31P[M] + 0.69(1-P[M])}, \text{ and} \\ P[M^c|A^c] &= \frac{P[A^c|M^c]P[M^c]}{P[A^c]} = 0.69 \frac{P[M^c]}{0.31P[M] + 0.69(1-P[M])}. \end{aligned}$$

As a partial check, we note that $P[M|A]+P[M^c|A] = 1$ as it must, and likewise for $P[M|A^c]+P[M^c|A^c]$. Then, for $P[M] = 10^{-3}$, we get $P[M|A] \simeq 2\times 10^{-3}$, $P[M^c|A] \simeq 0.998$, $P[M|A^c] \simeq 0.45 \times 10^{-3}$, and $P[M^c|A^c] \simeq 0.9996$. But, for $P[M] = 10^{-6}$, we get $P[M|A] \simeq 2.2 \times 10^{-6}$, $P[M^c|A] \simeq 0.999998$, $P[M|A^c] \simeq 0.45\times 10^{-36}$, and $P[M^c|A^c] \simeq 0.999998$. Thus, because of the uncertainty in the prior probability $P[M]$, these calcuated probability numbers have little value for decision making.

32.



A clearer version of the function is given below:

```
function [alpha,beta] = roc(r)
%function for evaluations in Problem 2.32
t=zeros (1,100); alpha=zeros(1,100); beta=zeros(1,100);
t(1)=-6;
for n=2:100
t(n)=t(1)+(1.5*n/100)*(r+6);
alpha(n)=quad('bell',t(n),10);
beta(n)=quad('bell',t(n)-r,10);
end
plot(alpha(2:100),beta(2:100))
axis([0 1 0 1])
axis ('image')
end
```

However, note that this function will only work with definition of the function 'bell', not given here. See documentation on MATLAB function 'quad'.

33. From the data, $\lambda = 9 \times 10^6$ ph/sec. For the counting interval (CI) $\Delta t = 10^{-6}$ sec. then,

$\lambda\Delta t = 9$. So

$$\begin{aligned} P[\{\text{false alarm in CI}\}] &= P[\{\text{ 0 photons in CI}\}] + P[\{\text{ 1 photon in CI}\}] \\ &= \frac{(9)^0}{0!}e^{-9} + \frac{(9)^1}{1!}e^{-9} \\ &= 10e^{-9} \simeq 0.0012. \end{aligned}$$

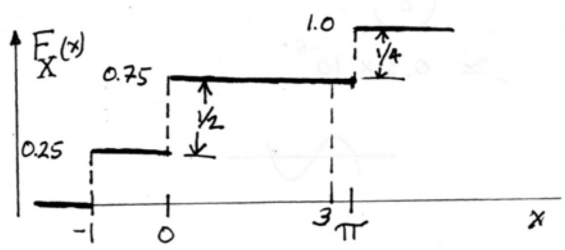
$$\begin{aligned} P[\{\text{at least one false alarm in } 10^6 \text{ tries}\}] &= 1 - P[\{\text{ 0 false alarms in } 10^6 \text{ tries}\}] \\ &\simeq 1 - \binom{10^6}{0}(0.0012)^0(1-0.0012)^{10^6} \\ &= 1 - (0.9988)^{10^6} \\ &\simeq 0. \end{aligned}$$

34. (a) $\Omega = \{G, R, Y\}$, where G=green, R=red, and Y=yellow. The σ-field of events are: $\{G\}, \{R\}, \{Y\}, \{G, R\}$(i.e. light is green or red) $, \{G, Y\}$(i.e.light is green or yellow), $\{R, Y\}$(light is red or yellow) $, \phi$(null event), and Ω, the certain event.

(b) $X(G) = -1, X(R) = 0, X(Y) = \pi$, and $P[G] = P[Y] = 0.5P[R]$. Hence

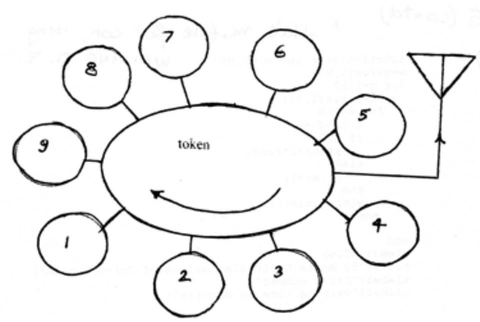
$$P[R] + 0.5P[R] + 0.5P[R] = 1.$$

So $P[R] = \frac{1}{2}$ and $P[G] = P[Y] = \frac{1}{4}$.



$F_X(3) = 0.75$

35.



(a)

$$\begin{aligned} T_{\max} &= 8 \times 5 = 40 \text{ minutes}, \\ T_{\min} &= 8 \times 0 = 0 \text{ minutes}. \end{aligned}$$

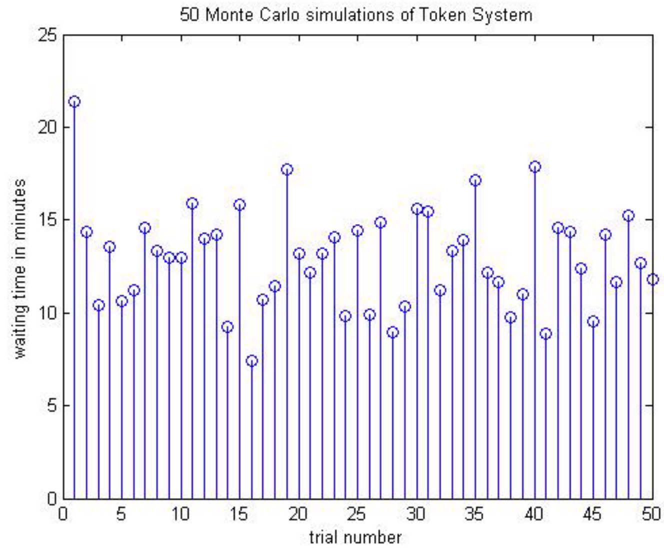
Note that if each station sends a message that is infinitesimally short, you get $T_{\min} = 0$. Let T denote the waiting time and let $p \triangleq P[\{\text{a station is busy}\}]$. Then

$$
\begin{aligned}
E[T] &= [2.5p + (1-p)1]\, 8 \\
&= 8 \text{ minutes for } p = 0, \\
&= 14 \text{ minutes for } p = 0.5, \\
&= 20 \text{ minutes for } p = 1.
\end{aligned}
$$

(b) Here is a MATLAB function that simulates the waiting time:

```
function [w]=token (p, NT, NS)
% function to simulate the token system in Problem 2.35
% p=probability a station is occupies, NT is number of trials,
% and NS is number of stations.
w=zeros (1, NT);
for n=1:NT
st=zeros(1,9);
for m=2:9
z=rand<=p;
if z>0
st(m)=5*rand;
else
st (m)=1;
end
w(n)=sum (st);
end
end
stem (1: NT, w)
title('50 Monte Carlo simulations of Token System')
xlabel('trial number')
ylabel('waiting time in minutes')
```

Here is a sample output, corresponding to 'probability of station occupied' $p = 0.4$, 'number of trials' $NT = 50$, and 'number of stations' $NS = 9$.



36.

$$\begin{aligned}
P[X^2+Y^2 \le c^2] &= \iint_{(x,y)\ :x^2+y^2\le c^2} \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} dxdy, \quad \text{transform to polar coordinates} \\
&= \frac{1}{2\pi}\int_0^c \int_0^{2\pi} e^{-\frac{1}{2}r^2} r dr d\theta, \quad \text{with } r=\sqrt{x^2+y^2} \text{ and } dxdy = rdrd\theta, \\
&= \int_0^c e^{-\frac{1}{2}r^2} rdr, \quad \text{let } u \triangleq \frac{1}{2}r^2, \quad \text{then} \quad du = rdr, \\
&= \int_0^{c^2/2} e^{-u} du \\
&= 1 - e^{-c^2/2} = 0.95.
\end{aligned}$$

Thus we need

$$\begin{aligned}
\frac{c^2}{2} &= \ln\frac{1}{1-0.95} = \ln 20 \simeq 3, \\
c &\simeq \sqrt{6} = 2.45.
\end{aligned}$$

37. (a) Since the area of this square with side $\sqrt{2}$ is 2, constant joint density $f_{X,Y}$ must take on value $\frac{1}{2}$ to be properly normalized, thus $A = \frac{1}{2}$.

(b) We can see four regions for the y values in evaluating

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy.$$

These regions are $x \le -1, -1 < x < 0, 0 \le x < 1$,and $x \ge 1$. Now, the first and last of these regions gives the trivial result $f_X(x) = 0$. For $0 \le x < 1$, we get

$$f_X(x) = \int_{x-1}^{1-x} \frac{1}{2}dy = \frac{1}{2}(1-x-x+1) = 1-x.$$

Similarly for $-1 < x < 0$, we get

$$f_X(x) = \int_{-x-1}^{1+x} \frac{1}{2} dy = \frac{1}{2}(1 + x + x + 1) = 1 + x.$$

Combining these regions we finally get

$$f_X(x) = \begin{cases} 1 - |x|, & |x| < 1, \\ 0, & \text{else.} \end{cases}$$

(c) If X is close to 1, then we see that Y must be close to 0. This suggests dependence between X and Y. To be sure we can use the result of part b together with the symmetry of the joint density to check whether $f_{X,Y} = f_X f_Y$ or not. By symmetry of $f_{X,Y}$ it must also be that

$$f_Y(y) = \begin{cases} 1 - |y|, & |y| < 1, \\ 0, & \text{else.} \end{cases}$$

Now the product of these two triangles $(1 - |x|)(1 - |y|) \neq \frac{1}{2}$ on $\text{supp}(f_{X,Y})$, so the random variables are definitely dependent. (The *support* of a function $f(x)$ is the set of domain values $\{x | f(x) \neq 0\}$ and is written as $\text{supp}\{f\}$.)

(d) We start with the definition and then plug in our result from part b:

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)} \\ &= \begin{cases} \frac{0.5}{1-|x|}, & 0 \leq |x| + |y| < 1, \\ 0, & \text{otherwise in } \{|x| < 1\}, \\ \times, & |x| \geq 1. \end{cases} \end{aligned}$$

Note that the conditional density is not defined for $\{|x| \geq 1\}$.

38. The pdf of the failure time random variable X is

$$\begin{aligned} f_X(t) &= \alpha(t) \exp\left(- \int_0^t \alpha(t') dt' \right) \\ &= \mu \exp(-\mu t) \text{ in this case.} \end{aligned}$$

Assume μ is measured in $(\text{hours})^{-1}$. If $A = \{\text{failure in 100 hrs or less}\}$, then

$$\begin{aligned} P[A] &= P[X \leq 100] \\ &= \int_0^{100} \mu e^{-\mu t} dt \\ &= 1 - e^{-\mu 100} \\ &\leq 0.05? \end{aligned}$$

Thus we need $e^{-\mu 100} \geq 0.95$, or taking logs and solving,

$$\mu \leq 5.13 \times 10^{-4}.$$

39. In general, for any $\alpha(t)$, the pdf of the failure time random variable X is

$$f_X(t) = \alpha(t) \exp\left(- \int_0^t \alpha(t')dt' \right).$$

(i) for $t < 0, \quad \alpha(t) = 0$, and so $f_X(t) = 0$,
(ii) for $0 \leq t \leq 10, \quad \alpha(t) = \frac{1}{2}$, and so

$$f_X(t) = \frac{1}{2}e^{-\frac{1}{2}t},$$

(iii) for $t > 10, \quad ,\alpha(t) = t - c$ for some constant c. Now at $t = 10, \alpha = \frac{1}{2}$, thus $\frac{1}{2} = 10 - c$, and so $c = 9.5$. Then, for $t > 10$,

$$\begin{aligned} f_X(t) &= (t - 9.5) \exp\left(- \left\{ \int_0^{10} \frac{1}{2} dt' + \int_{10}^t (t' - 9.5)dt' \right\} \right) \\ &= (t - 9.5) \exp\left(- \left\{ 5 + (\frac{1}{2}t^2 - 9.5t) - (50 - 95) \right\} \right) \\ &= (t - 9.5) \exp\left(- \left\{ \frac{t^2}{2} - 9.5t + 50 \right\} \right). \end{aligned}$$

We can put all this together in the one equation

$$f_X(t) = \begin{cases} 0, & t < 0, \\ \frac{1}{2}e^{-\frac{1}{2}t}, & 0 \leq t \leq 10 \\ (t - 9.5) \exp\left(- \left\{ \frac{t^2}{2} - 9.5t + 50 \right\} \right), & t > 10. \end{cases}$$

40. (a) Let $\Delta x > 0$ and $\Delta y > 0$, then

$$\begin{aligned} P[x &< X \leq x + \Delta x, y < Y \leq y + \Delta y] \\ &= F_{XY}(x + \Delta x, y + \Delta y) - F_{XY}(x + \Delta x, y) \\ &\quad - F_{XY}(x, y + \Delta y) + F_{XY}(x, y) \\ &= \left(\frac{F_{XY}(x + \Delta x, y + \Delta y) - F_{XY}(x, y + \Delta y)}{\Delta x} \right) \Delta x \\ &\quad - \left(\frac{F_{XY}(x + \Delta x, y) - F_{XY}(x, y)}{\Delta x} \right) \Delta x \\ &\simeq \left(\frac{\partial F_{XY}(x, y + \Delta y)}{\partial x} - \frac{\partial F_{XY}(x, y)}{\partial x} \right) \Delta x \\ &\simeq \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \Delta x \Delta y \\ &\simeq f_{XY}(x, y) \Delta x \Delta y. \end{aligned}$$

(b) By definition of the density f_{XY} as the mixed partial derivative of the distribution function F_{XY}, the integral of the density over all space must be $F_{XY}(\infty, \infty)$, since $F_{XY}(-\infty, \infty) = F_{XY}(\infty, -\infty) = F_{XY}(-\infty, -\infty)$ are all zero. But, again by definition,

$$\begin{aligned} F_{XY}(\infty, \infty) &= P[X \leq \infty, Y \leq \infty] \\ &= 1. \end{aligned}$$

(c) From part (a), it follows that

$$f_{XY}(x, y)\Delta x \Delta y \geq 0,$$

since it is a probability. Then since we took $\Delta x > 0$ and $\Delta y > 0$, it follows that

$$f_{XY}(x, y) \geq 0 \quad \text{too.}$$

Solutions to Chapter 3

1. We start with the definition of the CDF of Y,

$$\begin{aligned} F_Y(y) &\triangleq P[Y \leq y] \\ &= P[aX + b \leq y] \\ &= P[aX \leq y - b] \\ &= P\left[X \geq \frac{y-b}{a}\right], \qquad \text{for } a < 0. \end{aligned}$$

Now for any value x, we have $P[X \geq x] + P[X < x] = 1$ and $P[X < x] = F_X(x) - P[X = x]$. Hence for $a < 0$,

$$\begin{aligned} F_Y(y) &= P\left[X \geq \frac{y-b}{a}\right] \\ &= 1 - P\left[X < \frac{y-b}{a}\right] \\ &= 1 - \left(F_X(\frac{y-b}{a}) - P\left[X = \frac{y-b}{a}\right]\right) \\ &= 1 - F_X(\frac{y-b}{a}) + P\left[X = \frac{y-b}{a}\right]. \end{aligned}$$

2. We are given that X is a Gaussian random variable distributed as $N(0,1)$, i.e. $\mu = 0$ and $\sigma^2 = 1$. A second random variable is the result of the transformation $y = g(x)$, (Fig. 1)

$$g(x) = \begin{cases} x, & x \geq 0, \\ x^2, & x < 0. \end{cases}$$

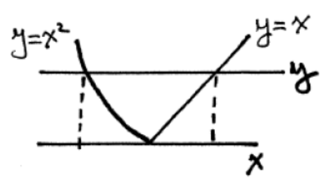


Figure 1:

We are asked to find the pdf of $Y = g(X)$. We proceed using the indirect method of first finding the CDF. From the figure we can see that, for $y > 0$, the event

$$\{Y \leq y\} = \{-\sqrt{y} \leq X \leq y\}.$$

For $y < 0$, the event $\{Y \leq y\} = \phi$, the null event, with probability zero. So the sought after CDF becomes

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= P[-\sqrt{y} \leq X \leq y] \\ &= F_X(y) - F_X(-\sqrt{y}). \end{aligned}$$

Taking the derivative with respect to y, we get

$$\begin{aligned} f_Y(y) &\triangleq \frac{dF_Y(y)}{dy} \\ &= \frac{dF_X(y)}{dy} - \frac{dF_X(-\sqrt{y})}{dy} \\ &= f_X(y) - f_X(-\sqrt{y})\frac{-d\sqrt{y}}{dy} \\ &= f_X(y) + f_X(-\sqrt{y})\frac{1}{2\sqrt{y}}. \end{aligned}$$

Now plugging in the standard Normal (Gaussian) distribution for X, we get

$$f_Y(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2} + \frac{1}{2\sqrt{2\pi y}}e^{-\frac{1}{2}y}, \qquad y > 0.$$

For $y < 0$, the density $f_Y(y) = 0$. Since X is a continuous random variable, there is no probability mass at $y = 0$, so we can set $f_Y(y) = 0$ there too. The overall pdf for Y then becomes

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2} + \frac{1}{2\sqrt{2\pi y}}e^{-\frac{1}{2}y}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

3. (*function of RV*) For a given value of $Y = y$, with $y > 0$, there are two solutions for x, i.e. $x_1(y) = \frac{1}{2}y$ and $x_2(y) = -y$ as shown in the Fig. 2.

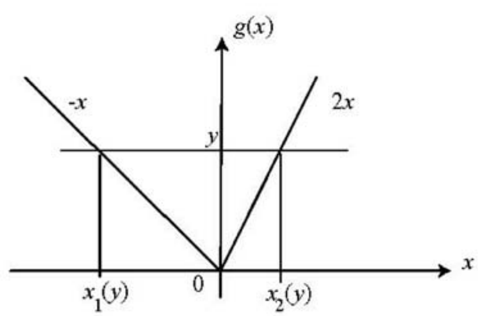


Figure 2:

Thus the pdf of Y is given as

$$\begin{aligned} f_Y(y) &= \left|\frac{dx_1(y)}{dy}\right| f_X(x_1(y)) + \left|\frac{dx_2(y)}{dy}\right| f_X(x_2(y)) \\ &= \left|\frac{1}{2}\right| f_X(y/2) + |-1|\, f_X(-y) \\ &= \frac{1}{2}f_X(y/2) + f_X(-y). \end{aligned}$$

When $Y < 0$, there are no solutions, so the general form of the pdf of Y is

$$f_Y(y) = \left(\frac{1}{2}f_X(y/2) + f_X(-y)\right)u(y),$$

where $u(y)$ is the unit step function. For this problem, we have $X \sim N(0,25)$,thus

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi 5}} e^{-\frac{1}{2}\frac{x^2}{25}} \\ &= \frac{1}{\sqrt{2\pi 5}} e^{-\frac{x^2}{50}}, \end{aligned}$$

so that

$$\begin{aligned} f_Y(y) &= \left(\frac{1}{2}\frac{1}{\sqrt{2\pi 5}} e^{-\frac{(y/2)^2}{50}} + \frac{1}{\sqrt{2\pi 5}} e^{-\frac{(-y)^2}{50}}\right) u(y) \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{10} e^{-\frac{y^2}{200}} + \frac{1}{5} e^{-\frac{y^2}{50}}\right) u(y). \end{aligned}$$

4. We are given that X is a Gaussian random variable distributed as $N(0,2)$, i.e. $\mu = 0$ and $\sigma^2 = 2$. A second random variable is the result of the transformation $y = g(x)$,

$$g(x) = \begin{cases} x, & x \geq 0, \\ 2x^2, & x < 0. \end{cases}$$

We are asked to find the probability density function (pdf) of $Y = g(X)$. We find the CDF first and then differentiate to find the pdf. From the equation for g we can see that, for $y > 0$, the event

$$\{Y \leq y\} = \{-\sqrt{\frac{y}{2}} \leq X \leq y\}.$$

For $y < 0$, the event $\{Y \leq y\} = \phi$, the null event, with probability zero. So the sought after CDF becomes

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= P[-\sqrt{\frac{y}{2}} \leq X \leq y] \\ &= F_X(y) - F_X(-\sqrt{\frac{y}{2}}). \end{aligned}$$

Taking the derivative with respect to y, we get

$$\begin{aligned} f_Y(y) &\triangleq \frac{dF_Y(y)}{dy} \\ &= \frac{dF_X(y)}{dy} - \frac{dF_X(-\sqrt{\frac{y}{2}})}{dy} \\ &= f_X(y) - f_X(-\sqrt{\frac{y}{2}}) \frac{-d\sqrt{\frac{y}{2}}}{dy} \\ &= f_X(y) + f_X(-\sqrt{\frac{y}{2}}) \frac{1}{4\sqrt{\frac{y}{2}}}. \end{aligned}$$

Now plugging in the standard Normal (Gaussian) distribution for X, we get

$$f_Y(y) = \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}y^2} + \frac{1}{8\sqrt{\pi \frac{y}{2}}} e^{-\frac{1}{8}y}, \qquad y > 0.$$

For $y < 0$, the density $f_Y(y) = 0$. Since X is a continuous random variable, there is no probability mass at $y = 0$, so we can set $f_Y(y) = 0$ there too. The overall pdf for Y then becomes

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}y^2} + \frac{1}{8\sqrt{\pi\frac{y}{2}}} e^{-\frac{1}{8}y}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

5. In both parts of this problem the random variable X has its probability density given as exponential with parameter $\alpha(> 0)$,i.e. $f_X(x) = \alpha e^{-\alpha x} u(x)$.

   (a) Here the function is given as $y = g(x) = x^3$, as shown in Fig. 3.

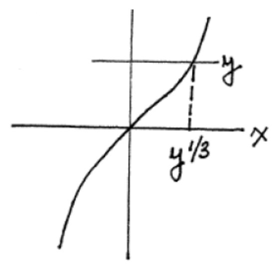


Figure 3:

Using the *indirect method*, we see that the distribution function of Y satisfies $F_Y(y) = F_X(y^{1/3})$ since the event $\{Y \leq y\} = \{X \leq y^{1/3}\}$. Upon differentiation we get

$$\begin{aligned} f_Y(y) &= dF_X(y^{1/3})/dy \\ &= f_X(y^{1/3}) \frac{dy^{1/3}}{dy} \\ &= \alpha e^{-\alpha y^{1/3}} u(y^{1/3}) \frac{1}{3} y^{-2/3} \\ &= \frac{\alpha}{3y^{2/3}} e^{-\alpha y^{1/3}} u(y). \end{aligned}$$

Alternatively, and since the function g is monotonic, we can easily use the *direct method* to get the pdf of $Y = g(X) = X^3$. We get

$$\begin{aligned} f_Y(y) &= f_X(y^{1/3}) \left|\frac{dx}{dy}\right| \\ &= f_X(y^{1/3}) / \left|\frac{dy}{dx}\right|. \end{aligned}$$

Now $\frac{dy}{dx} = 3x^2 = 3y^{2/3}$, so

$$\begin{aligned} f_Y(y) &= \alpha e^{-\alpha y^{1/3}} u(y^{1/3})/3y^{2/3} \\ &= \frac{\alpha}{3y^{2/3}} e^{-\alpha y^{1/3}} u(y), \quad \text{same as above.} \end{aligned}$$

(b) For this part, only the function g has changed, this time $g(x) = 2x + 3$. We choose to use the indirect method, and find

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= P\left[X \leq \frac{y-3}{2}\right] \\ &= F_X\left(\frac{y-3}{2}\right). \end{aligned}$$

Taking the derivative with respect to the free variable y, we get

$$f_Y(y) = \frac{1}{2} f_X\left(\frac{y-3}{2}\right),$$

which, for the exponential distribution with parameter α, becomes

$$\begin{aligned} f_Y(y) &= \frac{1}{2}\alpha e^{-\alpha\frac{y-3}{2}} u(\frac{y-3}{2}) \\ &= \frac{\alpha}{2} e^{-\frac{\alpha}{2}(y-3)} u(y-3). \end{aligned}$$

6. By definition $F_Y(y) = P[Y \leq y] = P[g(X) \leq y]$, so $F_Y(y) = 1$ for all $y \geq +1$ for this transformation g. When $-1 < y < +1$, we can write

$$\begin{aligned} F_Y(y) &= P[g(X) \leq y] \\ &= P[X \leq y] \\ &= F_X(y), \end{aligned}$$

then calculating the distribution function for the Laplacian density, we get by running integration

$$F_X(x) = \begin{cases} 1 - \frac{1}{2}e^{-x}, & x \geq 0, \\ \frac{1}{2}e^{+x}, & x < 0. \end{cases}$$

Thus, we have

$$\begin{aligned} F_Y(y) &= F_X(y), \qquad -1 < y < +1, \\ &= \begin{cases} 1 - \frac{1}{2}e^{-y}, & 1 > y \geq 0, \\ \frac{1}{2}e^{+y}, & -1 < y < 0. \end{cases} \end{aligned}$$

Clearly, for $y < -1$, we must have $F_Y(y) = 0$. Combining these results, we obtain

$$F_Y(y) = \begin{cases} 1, & y \geq 1, \\ 1 - \frac{1}{2}e^{-y}, & 1 > y \geq 0, \\ \frac{1}{2}e^{+y}, & 0 > y > -1, \\ 0, & -1 \geq y. \end{cases}$$

7. We present two methods to solve this transformation problem:

Method (1): For $y > 0$

$$P[Y \leq y] = P[e^X \leq y] = P[X \leq \ln y] = F_X(\ln y).$$

Hence

$$f_Y(y) = \frac{dF_X(\ln y)}{dy} = \frac{dF_X(\ln y)}{d(lny)}\frac{d(\ln y)}{dy} = \frac{1}{y}f_X(\ln y).$$

For $y \le 0$

$$P[Y \le y] = P[e^X \le y] = P[\phi] = 0.$$

Hence

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\ \sigma y}\exp[-\frac{1}{2}(\frac{\ln y - \mu}{\sigma})^2]u(y).$$

Method (2): A plot of $y = g(x)$ is is $x = \ln y$, for $y > 0$ as given in Fig. 4.

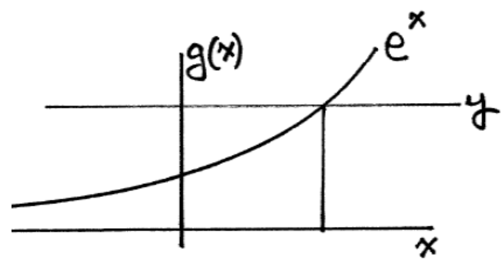


Figure 4:

$$f_Y(y) = \sum_{i=1}^{r}\frac{f_X(x_i)}{|g'(x_i)|} = f_X(\ln y)/e^x|_{x=\ln y}.$$

For $y \le 0$, there is no real solution to $y - g(x) = 0$; hence for $y \le 0$, the pdf of Y equals zero there. Combining solutions for both regions of y, we get

$$f_Y(y) = \frac{f_X(\ln y)}{y}u(y).$$

8. (a) Note that the peak is approximately at $\ln y = \mu$ or $y = \exp \mu$ when $\mu > \sigma$. To investigate the behavior near zero, to aid in a hand plot, we note

$$\begin{aligned}\frac{1}{y}e^{-(\ln y)^2} &= \frac{1}{y}e^{-(\ln y)(\ln y)} \\ &= \frac{1}{y}\left(e^{-\ln y}\right)^{\ln y} \\ &= \frac{1}{y}\left(\frac{1}{y}\right)^{\ln y} \\ &= y^{-\ln y - 1}.\end{aligned}$$

Now near zero, the exponent $-\ln y - 1$ takes on large positive values, thus $y^{-\ln y - 1}$ converges to zero as y approaches zero from the right. So the hand plots should just look to first order like a Gaussian density, with a log scale for the horizontal axis.

Precise plotting with $\mu = 2$ and $\sigma = 1$, results in the Fig. 5. The plots are in terms of conventional y versus x axis plots.

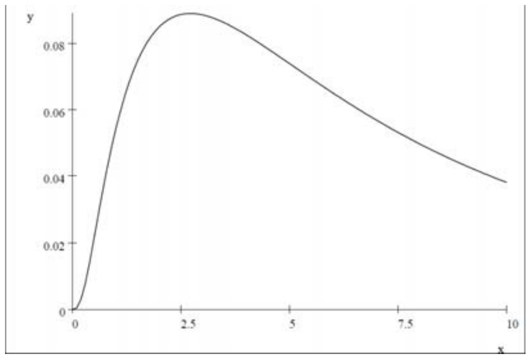


Figure 5:

$\frac{1}{\sqrt{2\pi}x} \exp -\frac{(\ln x-2)^2}{2}$
A second plot is with $\mu = 0$ and $\sigma = 4$ is given in Fig. 6.
$\frac{1}{\sqrt{2\pi}4x} \exp -\frac{(\ln x)^2}{32}$

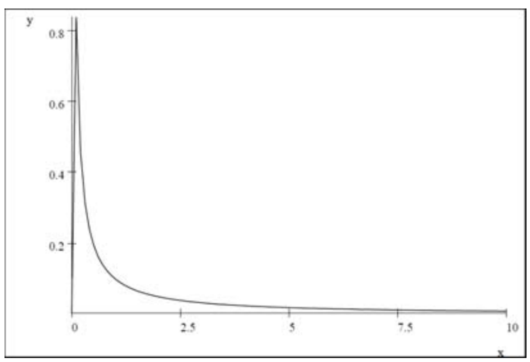


Figure 6:

(b) We first integrate the density $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma y} \exp -\frac{(\ln y-\mu)^2}{2\sigma^2}\ u(y)$. We have

$$F_Y(y) = 0 \text{ for } y < 0.$$

For $y \geq 0$, we write

$$F_Y(y) = \int_0^y \frac{1}{\sqrt{2\pi}\sigma s} \exp -\frac{(\ln s - \mu)^2}{2\sigma^2}\ ds.$$

We then make the substitution $t \triangleq \ln s$ to get (with $ds = e^t dt$)

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{\ln y} \frac{1}{\sqrt{2\pi}\sigma e^t} \exp -\frac{(t-\mu)^2}{2\sigma^2}\ e^t dt \\ &= \int_{-\infty}^{\ln y} \frac{1}{\sqrt{2\pi}\sigma} \exp -\frac{(t-\mu)^2}{2\sigma^2}\ dt \\ &= \frac{1}{2} + \operatorname{erf}\left(\frac{\ln y - \mu}{\sigma}\right). \end{aligned}$$

The overall solution for the distribution function (CDF) is then

$$F_Y(y) = \left[\frac{1}{2} + \text{erf}\left(\frac{\ln y - \mu}{\sigma}\right)\right] u(y).$$

The alternative solution is easier and starts with the distribution function

$$\begin{aligned}
F_Y(y) &\triangleq P[Y \leq y] \\
&= P[e^X \leq y] \\
&= P[X \leq \ln y], \quad \text{because } \ln(\cdot) \text{ is a monotonic increasing function,} \\
&= F_X(\ln y) \quad \text{on } y > 0, \\
&= \frac{1}{2} + \text{erf}\left(\frac{\ln y - \mu}{\sigma}\right), \quad \text{on } y > 0, \\
&= \left[\frac{1}{2} + \text{erf}\left(\frac{\ln y - \mu}{\sigma}\right)\right] u(y), \quad \text{since } F_Y(y) = 0 \text{ for } \quad y \leq 0.
\end{aligned}$$

9. *Method* (1):

$$P[Y \leq y] = P[\ln X \leq y] = P[X \leq e^y] = F_X(e^y)$$

$$f_Y(y) = \frac{dF}{d(e^y)}\frac{d(e^y)}{dy} = f_X(e^y)e^y = \frac{1}{3}e^{-\frac{1}{3}e^y}u(e^y)e^y$$

But since $e^y \geq 0$ all real y, $u(e^y) = 1$ everywhere. Hence

$$f_Y(y) = \frac{1}{3}e^{-\frac{1}{3}(\exp(y) - 3y)}$$

Method (2): We plot $y = g(x)$ first, in Fig. 7. The only solution to $y - g(x) = 0$ is $x = \ln y$

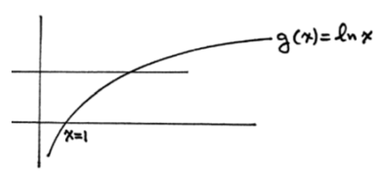


Figure 7:

for $-\infty < y < \infty$. Hence

$$\begin{aligned}
f_Y(y) &= \frac{f_X(e^y)}{|dg/dx|_{x=e^y}} \\
&= f_X(e^y)e^y.
\end{aligned} \tag{1}$$

10. Here is the plot of $y = g(x)$ in Fig. 8: $y - g(x) = 0$ is $y = \sqrt{x}$, or $x = y^2$. For $y < 0$, no real solutions exist to $y - g(x) = 0$. Hence

$$f_Y(y) = \frac{f_X(x)}{g'(x)}|_{x=y^2}.$$

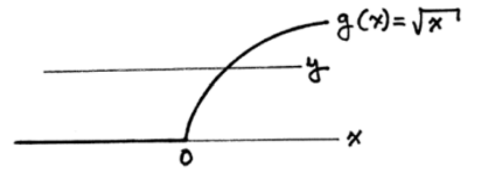


Figure 8:

Now

$$g'(x) = \frac{1}{2}x^{-\frac{1}{2}}|_{x=y^2} = \frac{1}{2y},$$

thus

$$f_Y(y) = \begin{cases} \sqrt{\frac{2}{\pi}}ye^{-\frac{1}{2}y^4}, & y > 0, \\ 0, & y < 0. \end{cases}$$

But at $y = 0$ there is a probability mass, jump in the distribution function, and impulse in the density function. We have

$$\begin{aligned} P[Y = 0] &= P[X \leq 0] \\ &= F_Y(0) = \frac{1}{2}, \end{aligned}$$

thus

$$F_Y(0) = \frac{1}{2}\int_{-\epsilon}^{\epsilon} \delta(y)dy.$$

So, combining results, we have the following answer valid for all y,

$$f_Y(y) = \frac{1}{2}\delta(y) + \sqrt{\frac{2}{\pi}}ye^{-\frac{1}{2}y^4}u(y).$$

11. (a) From the function g given in Fig. P3.11 in the text, we see that the event $\{Y \leq y\} = \{-\infty < X < +\infty\} = \Omega$[1] for $y \geq 1$. Hence $F_Y(y) = 1$ there. Next consider $0 < y < 1$, again from the figure, we see that in this region of the function g with slope 1, $\{Y \leq y\} = \{X \leq y\} \cup \{X > 2\}$, a disjoint union. So in this region of y values, it must be that $F_Y(y) = F_X(y) + (1 - F_X(2))$. Right at the point $y = 0$, we have $\{Y \leq y\} = \{X \leq 0\} \cup \{X > 2\}$. Now consider the remaining region $y < 0$, there the event $\{Y \leq y\} = \phi$ the null event, since there are no x values that map to $y < 0$ for the given function g shown in Fig. 9. So $F_Y(y) = 0$ there. Since $X : N(0, 1)$, we can write

$$\begin{aligned} F_Y(y) &= \begin{cases} 0, & y < 0, \\ \frac{1}{2} + \text{erf}(y) + \left[1 - \left(\frac{1}{2} + \text{erf}(2)\right)\right], & 0 \leq y < 1, \\ 1, & 1 \leq y, \end{cases} \\ &= \begin{cases} 0, & y < 0, \\ 1 + \text{erf}(y) - \text{erf}(2), & 0 \leq y < 1, \\ 1, & 1 \leq y. \end{cases} \end{aligned}$$

[1] Actually, we only said that random variables take on finite values with probability one. So there could be some events of probability zero that actually attain $X = \pm\infty$. However, since they are probability zero, the distribution function is not affected.

Taking derivatives with respect to the free variable y, we get *within* each interval

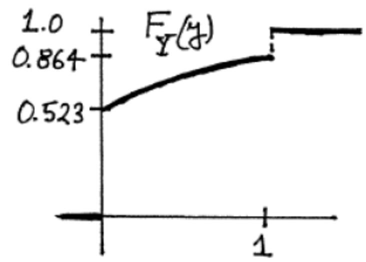


Figure 9:

$$f_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}, & 0 < y < 1, \\ 0, & 1 < y. \end{cases}$$

This is not the complete story though as there are jumps in the CDF $F_Y(y)$ at both $y = 0$ and $y = 1$. Thus we must add to this pdf two impulses located at these two jump points. From the CDF of Y given above, we see that

$$\begin{aligned} F_Y(1) - F_Y(1^-) &= 1 - (1 + \text{erf}(1) - \text{erf}(2)) \\ &= \text{erf}(2) - \text{erf}(1) \\ &\doteq 0.137 \end{aligned}$$

so this is the area of the impulse needed at $y = 1$. It thus becomes $0.137\delta(y-1)$. The jump in the CDF $F_Y(y)$ at $y = 0$ is found as

$$\begin{aligned} F_Y(0) - F_Y(0^-) &= (1 + \text{erf}(0) - \text{erf}(2)) - 0 \\ &= 1 - \text{erf}(2) \\ &\doteq 0.523, \end{aligned}$$

so the needed impulse in the density at $y = 0$ becomes $0.523\delta(y)$. Combining these results, we get the pdf $f_Y(y)$ over the full domain of y as

$$f_Y(y) = 0.523\delta(y) + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} (u(y) - u(y-1)) + 0.137\delta(y-1).$$

(b) Equation 3.2-23 in the book will only give us the portion of the solution for $0 < y < 1$, where in fact there is one root $x_1(y) = y$ and $n = 1$. Using this method, we thus get $f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$ and miss the impulses at $y = 0$ and $y = 1$. The problem here is that our function g has regions of zero slope. Over such regions, there are uncountably infinitely many x values corresponding to the one y value. Thus Equation 3.2-23 does not hold there.

12. The given pdf of X is uniform over $[0, 2]$. Thus $f_X(x) = \frac{1}{2}(u(x) - u(x-2))$. The function g (Fig. 10) is given as

$$g(x) = \begin{cases} 0, & x < 0, \\ 2x, & 0 \le x < \frac{1}{2}, \\ 2 - 2x, & \frac{1}{2} \le x < 1, \\ 0, & 1 \le x, \end{cases}$$

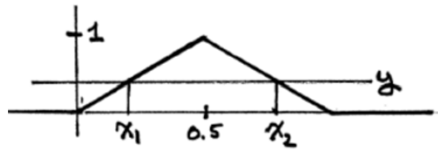


Figure 10:

For $y \geq 1$, for the given function g, the event $\{Y \leq y\} = \{-\infty < X < +\infty\}$, so for all $y \geq 1$, the CDF $F_Y(y) = 1$. When $0 < y \leq 1$, there are two solutions to the equation $y = g(x)$, and they can be given as $x_1(y) = y/2$ and $x_2(y) = 1 - y/2$. So, using Equation 3.2-23, we get

$$f_Y(y) = \frac{1}{2} f_X(y/2) + \frac{1}{2} f_X(1 - y/2), \qquad 0 < y \leq 1.$$

At $y = 0$, we see the probability mass $P[Y = 0] = P[X \leq 0] + P[X \geq 1]$, so we must add an impulse of this area to the density $f_Y(y)$ at $y = 0$. The overall answer for the general pdf f_Y then becomes,

$$f_Y(y) = \left(\frac{1}{2} f_X(y/2) + \frac{1}{2} f_X(1 - y/2)\right)(u(y) - u(y-1)) + (P[X \leq 0] + P[X \geq 1])\,\delta(y).$$

For the given uniform density of random variable X, i.e. $X : U[0, 2]$, we have $P[X \leq 0] = 0$ and $P[X \geq 1] = 1/2$. Also

$$f_X(y/2) = \frac{1}{2}\left(u(y/2) - u(y/2 - 2)\right) = \frac{1}{2}\left(u(y) - u(y-4)\right)$$

and

$$f_X(1 - y/2) = \frac{1}{2}\left(u(1 - y/2) - u(1 - y/2 - 2)\right) = \frac{1}{2}\left(u(2 - y) - u(-y - 2)\right)$$

so that

$$f_Y(y) = \left(\frac{1}{4}\left(u(y) - u(y-4)\right) + \frac{1}{4}\left(u(2-y) - u(-y-2)\right)\right) \times \left(u(y) - u(y-1)\right) + \frac{1}{2}\delta(y).$$

If we define a function

$$\mathrm{rect}(x) \triangleq \begin{cases} 1, & -\frac{1}{2} < x < +\frac{1}{2}, \\ 0, & \text{else}, \end{cases}$$

then we can write $f_X(x) = \frac{1}{2}\mathrm{rect}(\frac{x-1}{2})$ and so $f_X(y/2) = \frac{1}{2}\mathrm{rect}\left(\frac{y/2-1}{2}\right) = \frac{1}{2}\mathrm{rect}(\frac{y-2}{4})$ and $f_X(1 - y/2) = \frac{1}{2}\mathrm{rect}\left(\frac{1-y/2-1}{2}\right) = \frac{1}{2}\mathrm{rect}(\frac{-y}{4}) = \frac{1}{2}\mathrm{rect}(\frac{y}{4})$. The overall answer for the density f_Y can then be written as

$$\begin{aligned} f_Y(y) &= \left(\frac{1}{4}\mathrm{rect}(\frac{y-2}{4}) + \frac{1}{4}\mathrm{rect}\left(\frac{y}{4}\right)\right)\mathrm{rect}(y - \frac{1}{2}) + \frac{1}{2}\delta(y), \\ &= \frac{1}{2}\mathrm{rect}(y - \frac{1}{2}) + \frac{1}{2}\delta(y). \end{aligned}$$

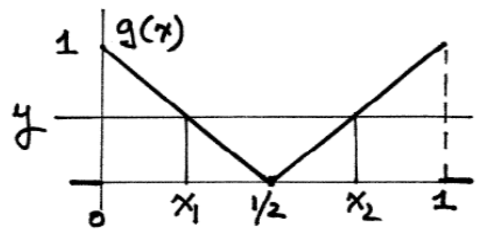


Figure 11:

13. First we plot $y = g(x)$ in Fig. 11. For $0 < x < \frac{1}{2}$: $g(x) = 1 - 2x$. For $\frac{1}{2} < x < 1$: $g(x) = 2x - 1$; elsewhere $g(x) = 0$. For $y > 1$, no real roots to $y - g(x) = 0$, then pdf $= 0$ there. For $y = 1$, $P[Y = 1] = P[X = 0] + P[X = 1] = 0$ since X is a continuous random variable. For $y < 0$, there are no real roots to $y - g(x) = 0$, then pdf=0. For $0 < y < 1$, there are two roots: $x_1 = \frac{1}{2} - \frac{y}{2}$; $x_2 = \frac{1}{2} + \frac{y}{2}$. So

$$f_Y(y) = \sum_{i=1}^{N} \frac{f_X(x_i)}{|dg/dx|_{x_i}} = \frac{1}{2} f_X(\frac{1}{2} - \frac{y}{2}) + \frac{1}{2} f_X(\frac{1}{2} + \frac{y}{2}).$$

For $y = 0$, since $X : U(0, 2)$

$$P[Y = 0] = P[X < 0] + P[X > 1] = \frac{1}{2}.$$

Therefore at $y = 0$, $f_Y(y) = \frac{1}{2}\delta(y)$. To construct $f_Y(y)$:

$$f_X(x) = \frac{1}{2}\text{rect}\left(\frac{x-1}{2}\right).$$

Therefore for $0 < y < 1$:

$$f_Y(y) = \frac{1}{4}\text{rect}\left(\frac{-y-1}{4}\right) + \frac{1}{4}\text{rect}\left(\frac{y-1}{4}\right).$$

And $\text{rect}(x) = \text{rect}(-x)$ in general. Therefore

$$f_Y(y) = \frac{1}{4}\text{rect}\left(\frac{y+1}{4}\right) + \frac{1}{4}\text{rect}\left(\frac{y-1}{4}\right).$$

for $0 < y < 1$, with sketch (Fig. 12) valid in $(0, 1)$, For all y, we thus have,

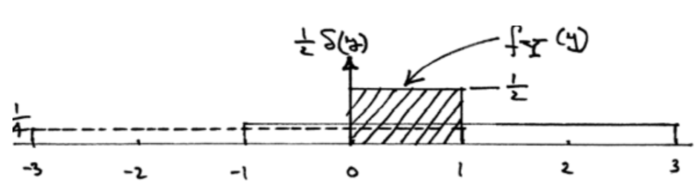


Figure 12:

$$f_Y(y) = \begin{cases} \frac{1}{2}\mathrm{rect}(y - \frac{1}{2}), & 0 < y < 1, \\ \frac{1}{2}\delta(y), & y = 0, \\ 0, & y < 0 \quad \text{or} \quad y > 1. \end{cases}$$

14. (a) See the sketch in Fig. 13.

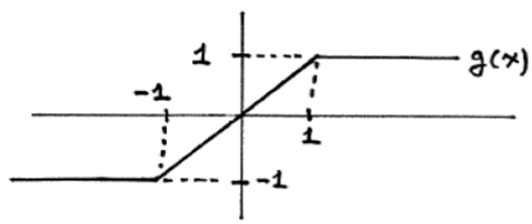


Figure 13:

(b) $F_Y(y)$ can be computed indirectly without computing $f_Y(y)$, but it is easier to do part (c) first and then come back to (b)

(c) For $y < -1$: no real solutions to $g(x) - y = 0$. For $y = -1$: $P[Y = -1] = P[X \leq -1] = P[X > 1] = \frac{1}{2} - \mathrm{erf}(1) \doteq 0.159$, so $f_Y(y) = 0.159\delta(y + 1)$. For $-1 < y < 1$: $g(x) = x$, $y - x = 0$, or $x = y$. Therefore $f_Y(y) = f_X(y)$. For $y = 1$: $P[Y = 1] = P[X > 1] \simeq 0.159$, so $f_Y(y) = 0.159\delta(y - 1)$. For $y > 1$: no real solutions to $g(x) - y = 0$ exist. Putting all these pieces together we get:

$$f_Y(y) = 0.159\delta(y + 1) + \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\mathrm{rect}(\frac{y}{2}) + 0.159\delta(y - 1),$$

which is sketched in Fig. 14. (b) Now that we have $f_Y(y)$, we obtain $F_Y(y)$ from

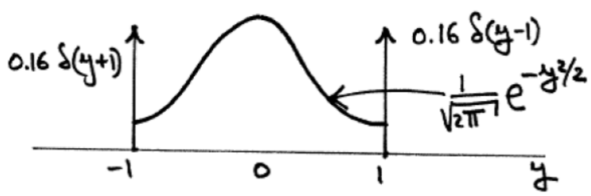


Figure 14:

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} f_Y(u)du \\ &= 0.159u(y + 1) + 0.159u(y - 1) \\ &\quad + \begin{cases} \mathrm{erf}(1) - \mathrm{erf}(-y), & -1 < y < 0, \\ \mathrm{erf}(1) + \mathrm{erf}(y), & 0 < y < 1. \end{cases} \end{aligned}$$

Note that this last part, involving the erf function is to be added to the step functions on the line above. Another way of writing the solution comes from noticing that for

$-1 < y < +1$, the distribution function of Y must agree with that of X. Thus in this region, we have the total answer $F_Y(y) = F_X(y) = 0.5+\text{erf}(y)$. Then we can write the solution as

$$F_Y(y) = \begin{cases} 1, & y \geq 1, \\ 0.5 + \text{erf}(y), & -1 \leq y < +1, \\ 0, & y < -1, \end{cases}$$

which is the same as that a few lines above. A sketch of the distribution function is provided in Fig. 15.

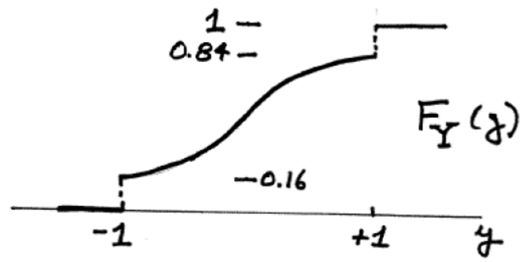


Figure 15:

15.

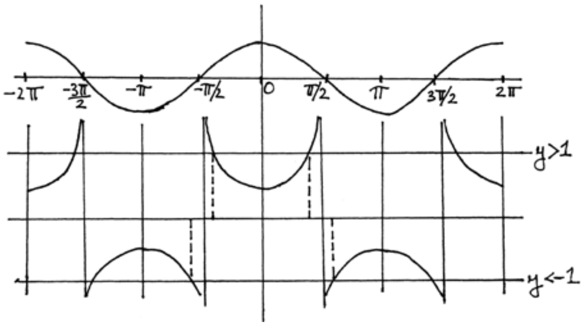
$$\begin{aligned} F_Y(y) &\triangleq P[Y \leq y] \\ &= P[\frac{a}{X} \leq y] \\ &= P[X \geq \frac{a}{y}] \\ &= 1 - P[X < \frac{a}{y}] \\ &= 1 - P[X \leq \frac{a}{y}], \text{ since } X \text{ is continuous,} \\ &= 1 - F_X(\frac{a}{y}). \end{aligned}$$

So, upon differentiation

$$\begin{aligned} f_Y(y) &= \frac{a}{y^2} f_X(\frac{a}{y}) \\ &= \frac{a}{y^2} \frac{\alpha/\pi}{\alpha^2 + (\frac{a}{y})^2} \\ &= \frac{a}{\alpha\pi} \frac{1}{y^2 + (a/\alpha)^2}. \end{aligned}$$

16.

$$\begin{aligned} Y &= \sec X \\ &= \frac{1}{\cos X} \end{aligned}$$



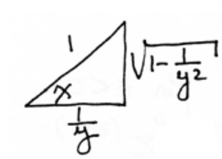
For $y > 1$, there are two roots in the interval $(-\pi, +\pi)$

$$\begin{aligned} y &= g(x) \\ &= \frac{1}{\cos x} \\ \Longrightarrow & \quad x = \cos^{-1}(1/y). \end{aligned}$$

The two roots are

$$x_1 = +\cos^{-1}(1/y) \quad \text{and} \quad x_2 = -\cos^{-1}(1/y).$$

Also $g(x) = (\cos x)^{-1} \quad \Longrightarrow \quad g'(x) = \frac{\sin x}{\cos^2 x}$. Can then use the right triangle



to show that $|g'(x_i(y))| = y\sqrt{y^2 - 1}$. Hence

$$\begin{aligned} f_Y(y) &= \sum_{i=1}^{2} \frac{f_X(x_i(y))}{|g'(x_i(y))|} \\ &= \frac{1}{\pi|y|} \frac{1}{y^2 - 1}. \end{aligned}$$

For $y < -1$, there are again two roots and we actually obtain the same result as above.

For $-1 < y < +1$, there are no solutions to $y - g(x) = 0$, thus $f_Y(y) = 0$ there.

Overall we have the solution

$$f_Y(y) = \begin{cases} \frac{1}{\pi|y|}\frac{1}{y^2-1}, & |y| > 1, \\ 0, & \text{else.} \end{cases}$$

17. Since the random variables X and Y are independent, we can find the density of Z via convolution of the two uniform densities f_X and f_Y, thus

$$\begin{aligned} f_Z(z) &= f_X(z) * f_Y(z) \\ &= \int_{-\infty}^{+\infty} f_X(z - x) f_Y(x) dx. \end{aligned}$$

Here, by the problem statement $X : U(-1, +1)$ and hence

$$f_X(x) = \begin{cases} \frac{1}{2}, & -1 < x < +1, \\ 0, & \text{else.} \end{cases}$$

Similarly $Y : U(-2, +2)$ and so

$$f_Y(y) = \begin{cases} \frac{1}{4}, & -2 < y < +2, \\ 0, & \text{else.} \end{cases}$$

Computing the convolution graphically, we see that the resulting function f_Z will be constant when the short pulse f_X is completely contained inside the longer pulse f_Y, and that this will occur for $-1 \leq z \leq +1$, for which the area is easily computed as $\frac{1}{2} \times \frac{1}{4} \times 2 = \frac{1}{4}$. From graphical considerations, we can also easily see that the output function f_Z must be zero when the two pulses do not overlap, and that this will occur for all $|z| > 3$. For $3 \geq |z| > 1$, we then just connect these result together via straight lines, to obtain

$$f_Z(z) = \begin{cases} 0, & z < -3, \\ \frac{1}{8}(z+3), & -3 \leq z < -1 \\ \frac{1}{4}, & -1 \leq z \leq +1, \\ \frac{1}{8}(-z+3), & +1 < z \leq +3 \\ 0, & z > +3 \end{cases},$$

which is graphed as Fig. 16.

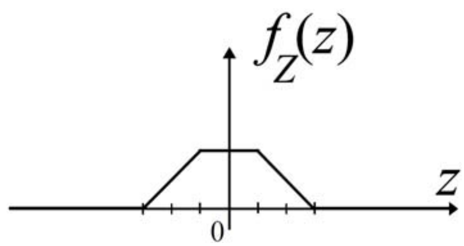


Figure 16:

18. We wish to calculate the pdf of $Z \triangleq X - Y$, given that X and Y are independent and exponentially distributed as $f_X(x) = f_Y(x) = \alpha \exp(-\alpha x) u(x)$. Our approach is to first find the density of $V \triangleq -Y$ and then make use of the convolution of densities property to find the density of $Z = X + V$. Now $P[V \leq v] = P[Y \geq -v] = 1 - F_Y(-v)$, where we have made use of the fact that Y is a continuous random variable. Therefore

$$\begin{aligned} f_V(v) &\triangleq \frac{dF_V(v)}{dv} \\ &= \frac{d(1 - F_Y(-v))}{dv} \\ &= f_Y(-v). \end{aligned}$$

Then, since X and V are independent,

$$\begin{aligned}
f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_V(z-x)dx \\
&= \int_{-\infty}^{+\infty} f_X(x) f_Y(-(z-x))dx \\
&= \int_{-\infty}^{+\infty} f_X(x) f_Y(x-z)dx \\
&= \int_{-\infty}^{+\infty} \alpha e^{-\alpha x} u(x) \alpha e^{-\alpha(x-z)} u(x-z)dx \\
&= \int_{\max(0,z)}^{+\infty} \alpha e^{-\alpha x} \alpha e^{-\alpha(x-z)} dx \\
&= \int_{\max(0,z)}^{+\infty} \alpha e^{-\alpha x} \alpha e^{-\alpha(x-z)} dx \\
&= \alpha^2 e^{+\alpha z} \int_{\max(0,z)}^{+\infty} e^{-2\alpha x} dx \\
&= \alpha^2 e^{+\alpha z} \begin{cases} \frac{1}{2\alpha}, & z < 0, \\ \frac{1}{2\alpha} e^{-2\alpha z}, & z \geq 0, \end{cases} \\
&= \frac{\alpha}{2} \begin{cases} e^{+\alpha z}, & z < 0, \\ e^{-\alpha z}, & z \geq 0, \end{cases} \\
&= \frac{\alpha}{2} \exp(-\alpha|z|), \qquad -\infty < z < +\infty.
\end{aligned}$$

19. (a) The joint pdf $f_{V,W}(v,w)$ is given as

$$f_{V,W}(v,w) = f_{X,Y}\left(\frac{1}{3}(v+w), \frac{1}{3}(2v-w)\right) \left|\frac{\partial(\phi,\psi)}{\partial(v,w)}\right|$$

with the transformation and inverse given as

$$\begin{array}{ll} v = x + y & x = \phi(v,w) = \frac{1}{3}(v+w) \\ w = 2x - y & y = \psi(v,w) = \frac{1}{3}(2v-w). \end{array}$$

We proceed to evaluate the absolute value of the Jacobian as

$$|J| = \left|\det\begin{bmatrix} \frac{\partial\phi}{\partial v} & \frac{\partial\phi}{\partial w} \\ \frac{\partial\psi}{\partial v} & \frac{\partial\psi}{\partial w} \end{bmatrix}\right| = \left|\det\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}\right| = \left|-\frac{1}{9} - \frac{2}{9}\right| = \frac{1}{3}.$$

Thus we finally obtain

$$f_{V,W}(v,w) = \frac{1}{3} f_{X,Y}\left(\frac{1}{3}(v+w), \frac{1}{3}(2v-w)\right).$$

(b) We start by noting that Z is just the V of part (a). Hence $f_Z(z)$ is just the marginal density for $f_{Z,W}(z,w)$ where

$$f_{Z,W}(z,w) = \frac{1}{3} f_{X,Y}\left(\frac{1}{3}(z+w), \frac{1}{3}(2z-w)\right).$$

Combining we have

$$f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{3} f_{X,Y}\left(\frac{1}{3}(z+w), \frac{1}{3}(2z-w)\right) dw.$$

Upon the substitution $x = \frac{1}{3}(z+w)$ inside the integral, we get

$$dx = \frac{1}{3} dw \quad \text{and} \quad \frac{1}{3}(2z-w) = \frac{1}{3}(2z-(3x-z)) = z-x,$$

so that the integral expression then becomes

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x, z-x)\, dx.$$

To get the desired result, we need that X and Y be independent RVs, because then the joint density $f_{X,Y}(x, z-x)$ will factor into $f_X(x) f_Y(z-x)$, which is the desired integrand.

20. In this problem

$$g(x) = \begin{cases} x, & |x| \leq 1, \\ 0, & \text{else,} \end{cases}$$

so the range of y is $0 \leq y \leq 1$. For $0 < y \leq 1$, i.e. when the first inequality is strict, there is only one root root to $y - g(x) = y - x = 0$: at $x_1 = y$. Hence

$$f_Y(y) = \frac{f_X(x_1)}{\left|\frac{dg}{dx}\right|} = \frac{f_X(x_1)}{1} \quad \text{for } 0 < y \leq 1.$$

Since $f_X(x) = e^{-x} u(x)$, it follows that

$$\begin{aligned} f_Y(y) &= f_X(x_1) \\ &= e^{-y} u(y) \\ &= e^{-y} \quad \text{there, i.e. for } 0 < y \leq 1. \end{aligned}$$

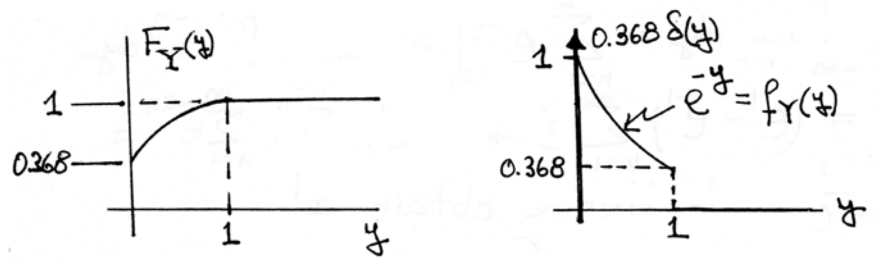
However, according to the mapping $g(x)$, the event $\{Y = 0\} = \{X \leq 0\} \cup \{X > 1\}$, so

$$\begin{aligned} P[Y=0] &= P[X \leq 0] + P[X > 1] \\ &= 0 + \int_1^{\infty} e^{-x} dx \\ &= e^{-1}. \end{aligned}$$

Thus the complete answer for the pdf of Y is:

$$\begin{aligned} f_Y(y) &= e^{-1}\delta(y) + e^{-y}[u(y) - u(y-1)] \\ &\simeq 0.368\delta(y) + e^{-y}[u(y) - u(y-1)]. \end{aligned}$$

Labeled sketches of the CDF and pdf follow:



21. The only real roots occur when $-1 < y < +1$, and, as can be seen from the diagram, these occur at $x_n = y + 2n$ for $n = ..., -2, -1, 0, 1, 2, ...$ Since $f_X(x) = e^{-x}u(x)$ and $|g'(x))|_{x=x_n} = 1$, we obtain

$$f_Y(y) = \sum_{n=-\infty}^{+\infty} e^{-(y+2n)}\text{rect}(y/2)u(y+2n),$$

where $\text{rect}(x) \triangleq u(x+\frac{1}{2})-u(x-\frac{1}{2}))$. We note that the product $\text{rect}(y/2)u(y+2n)$ =$\text{rect}(y/2)$ for $n \geq 1$. For $n < 0$, this product is zero. At $n = 0$, only half of $\text{rect}(y/2)$ is overlapped by $u(y)$. Hence, for $n = 0$, the product is $\text{rect}(y - \frac{1}{2})$. Thus,

$$f_Y(y) = \sum_{n=1}^{\infty} e^{-(y+2n)}\text{rect}(y/2) + e^{-y}\text{rect}(y - \frac{1}{2}).$$

To show that the pdf f_Y is a legitimate pdf, we note first that it is non-negative, and next that

$$\begin{aligned}
\int_{-\infty}^{+\infty} f_Y(y)dy &= \sum_{n=1}^{\infty} e^{-2n} \int_{-1}^{+1} e^{-y}dy + \int_{0}^{+1} e^{-y}dy \\
&= \left(\sum_{n=1}^{\infty} e^{-2n}\right)(e - e^{-1}) + (1 - e^{-1}) \\
&= \left(\frac{1}{1-e^{-2}} - 1\right)(e - e^{-1}) + (1 - e^{-1}) \\
&= \left(\frac{e}{e-e^{-1}} - 1\right)(e - e^{-1}) + (1 - e^{-1}) \\
&= e - (e - e^{-1}) + (1 - e^{-1}) \\
&= 1 \quad \text{as required for a legitimate pdf.}
\end{aligned}$$

22. Let X_n and X_{n+1} be two numbers generated in sequence by the random number generator at the nth construct. Form $Y_n = X_n + X_{n+1}$. If the $\{X_i\}$ are generated independently, the pdf of Y_n is $f_{Y_n}(y) = f_{X_n}(y) * f_{X_{n+1}}(y)$ and looks like the sketch in Fig. 17. So the form is right but the mean is wrong! Therefore generate $Z_n = Y_n - 1$. The pdf of Z_n looks like the sketch in Fig. 18. Equivalently one could subtract 0.5 from X_n and X_{n+1} before adding. In either case one can get $f_{Z_n}(z)$ =$\text{tri}(z/2)$. Here the *triangle function* tri is defined as

$$\text{tri}(x) \triangleq \begin{cases} 1 - 2|x|, & -\frac{1}{2} < x < +\frac{1}{2}, \\ 0, & \text{else.} \end{cases}$$

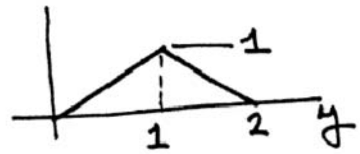


Figure 17:

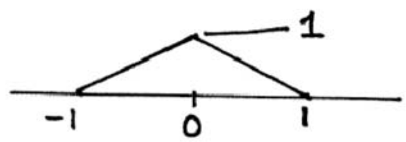
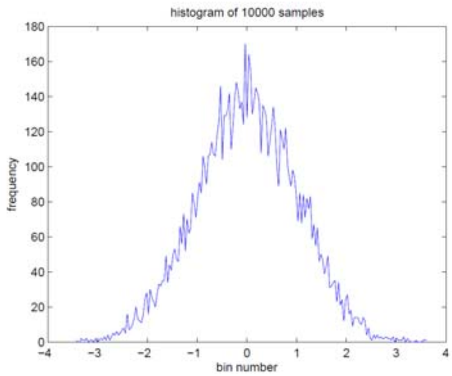


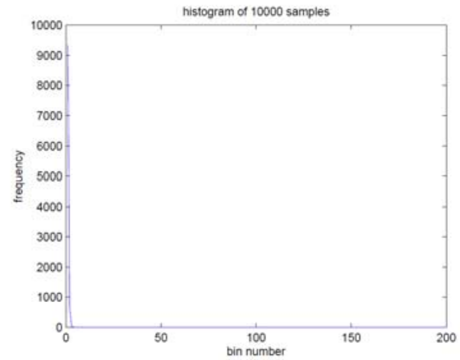
Figure 18:

23. We present a MATLAB function `histonorm` as follows:

```
function [alpha] = histonorm( samples, sumrange, bins );
sum = zeros( 1, samples );
for i = 1 : bins
x(i) = i;
end
for i = 1 : samples
for j = 1 : sumrange
sum(i) = sum(i) + rand;
end
sum(i) = sum(i) - 6;
end
[z, x] = hist( sum, bins );
plot( x, z);
title( ['histogram of ', num2str(samples), ' samples'] );
xlabel( 'bin number' )
ylabel( 'frequency' )
```

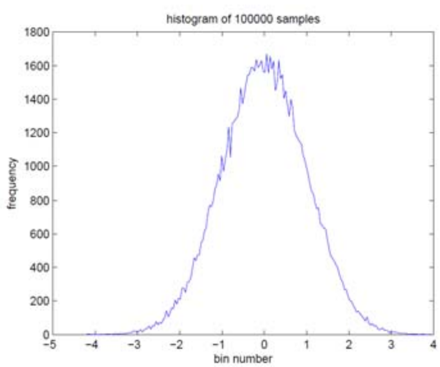
The multiple figures in Fig. **??** show the output of `histonorm` for 10,000 with "smart binning" on the left and "dumb binning" on the right.

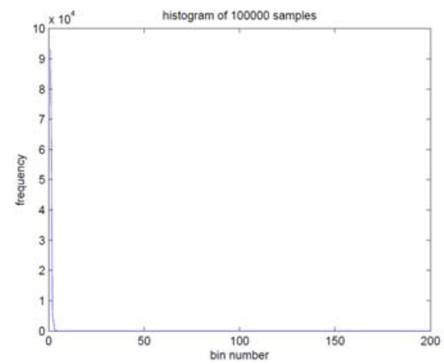


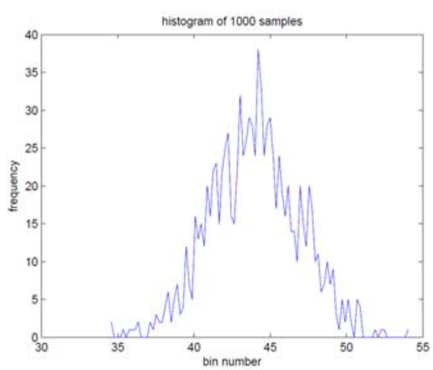


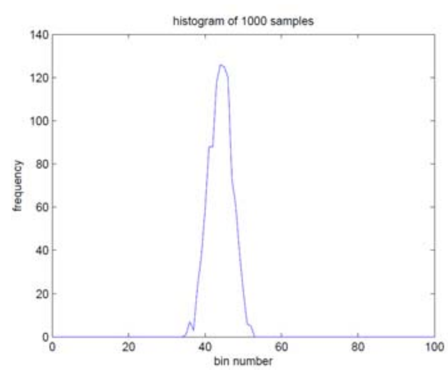
Note that the 'dumb binning' result on the right requires a slight modification of the function as indicated in the comments inside it. The 'smart binning' result which is the default only plots the domain where the bins are not empty.

While only 10,000 runs were requested in the problem statement, we next show in Fig. **??** the result of a larger number of trials 100,000 (Fig. **??** and **??**), and also a smaller number 1000 trials (Fig. **??** and **??**). We can see that the histogram begins to look strongly Gaussian at the high value of 100,000 trials. A better idea of its Gaussian-ness could be obtained by viewing a log plot, which is not done here however.









24. (a) We have to calculate the running integral of the density $f_Y(y) = \frac{c}{2}\exp(-c|y|)$. Now

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} f_Y(u)du \\ &= \int_{-\infty}^{y} \frac{c}{2}\exp(-c|u|)du. \end{aligned}$$

Because of the absolute value sign, it is easier to consider the two cases $y \leq 0$ and $y \geq 0$ separately. First we evaluate for $y \leq 0$, where $f_Y(y) = \frac{c}{2}\exp(+cy)$. We find

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} \frac{c}{2}\exp(cu)du \\ &= \frac{c}{2}\left(\frac{1}{c}\exp(cu)|_{-\infty}^{y}\right) \\ &= \frac{1}{2}\exp\left(cy\right), \qquad \text{for } y \leq 0. \end{aligned}$$

Now we consider the case $y \geq 0$, where $f_Y(y) = \frac{c}{2}\exp(-cy)$. We note that by symmetry $\int_{-\infty}^{0} \frac{c}{2}\exp(cu)du = 1/2$, so we can write

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} \frac{c}{2}\exp(-c|u|)du \\ &= \frac{1}{2} + \int_{0}^{y} \frac{c}{2}\exp(-cu)du \\ &= \frac{1}{2} + \frac{c}{2}\left(\frac{1}{c}\exp(-cu)|_{0}^{y}\right) \\ &= \frac{1}{2} + \frac{1}{2}\left(1-\exp\left(-cy\right)\right) \\ &= 1 - \frac{1}{2}\exp\left(-cy\right), \qquad \text{for } \ y \geq 0. \end{aligned}$$

Now, as a check, we note that both results agree at their common point $y = 0$ as they should. Overall, we can write the Laplacian distribution function as

$$F_Y(y) = \begin{cases} \frac{1}{2}\exp\left(cy\right), & y < 0, \\ 1 - \frac{1}{2}\exp\left(-cy\right), & y \geq 0. \end{cases}$$

(b) Probably the first thing to do here is to note that since $X : U[0,1]$, we have that $F_X(x) = x\{u(x) - u(x-1)\}$, i.e. just a straight line segment with slope 1 on $[0,1]$. Since the distribution function F_Y is monotone increasing, we have

$$F_Z(z) = P[X \leq g^{-1}(z)]$$

Next, we note that since $g = F_Y^{-1}$, so $g^{-1} = F_Y$ and hence

$$\begin{aligned} F_Z(z) &= P[X \leq g^{-1}(z)] \\ &= F_X(F_Y(z)) \\ &= F_Y(z)\left(u(F_Y(z)) - u(F_Y(z)-1)\right) \\ &= F_Y(z)\left(1-0\right) \\ &= F_Y(z). \end{aligned}$$

Note: Although it was not asked for in this problem, to actually use this method on a computer, we would need to calculate $g(x)$. It is given on $(0,1)$ as

$$g(x) = \begin{cases} \frac{1}{c}\ln 2x, & 0 < x < \frac{1}{2}, \\ \frac{1}{c}\ln\frac{1}{2(1-x)}, & \frac{1}{2} \leq x < 1. \end{cases}$$

(c) This method strictly speaking will not work with either jumps or flat regions in the desired distribution function F_Y. In a flat region of F_Y, the corresponding g would be a vertical line, not be a valid function! At a jump of F_Y, $g = F_Y^{-1}$ will not be defined for some of the input values. This won't work either. On the other hand, there are simple modifications of this method that can get around these problems and make the basic method useful in both cases. One simply has to remove the flat regions from F_Y before finding the inverse function. At the jumps, where the inverse function would have a gap, just fill it in with a horizontal line. With these changes the basic method extends to both mixed and discrete distribution functions.

25. First we solve $f_Z(z)$ is as follows. Let $Z = X - Y$, and $f_X(x) = f_Y(x) = \alpha e^{-\alpha x} u(x)$. Let $V \triangleq -Y$, then $P[V \leq v] = P[Y \geq -v] = 1 - F_Y(-v)$, and so $f_V(v) = \frac{dF_V}{dv} = f_Y(-v)$. Therefore

$$f_V(v) = \alpha e^{\alpha v} u(-v).$$

Now we can write $Z = X + V$, so that we have the density of Z given as the convolution

$$f_Z(z) = \int_{-\infty}^{\infty} f_V(v) f_X(z-v) dv.$$

Evaluating for $z < 0$ (Fig. 19), we get

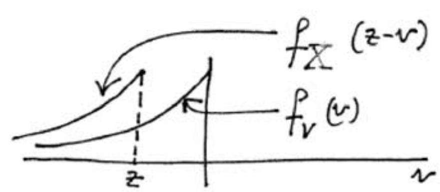


Figure 19:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_V(v) f_X(z-v) dv. \\ &= \int_{-\infty}^{z} e^{2\alpha v} dv \cdot e^{-\alpha z} \alpha^2 \\ &= \frac{\alpha}{2} e^{\alpha z}. \end{aligned}$$

For $z \geq 0$, we get

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{0} e^{2\alpha v} dv \cdot e^{-\alpha z} \alpha^2 \\ &= \frac{\alpha}{2} e^{-\alpha z}. \end{aligned}$$

Combining these two results, we get the formula valid for all z,

$$f_Z(z) = \frac{\alpha}{2} e^{-\alpha |z|},$$

as found in problem 3.18. Now let

$$W \triangleq |Z| = \begin{cases} Z, & Z \geq 0; \\ -Z, & Z < 0. \end{cases}$$

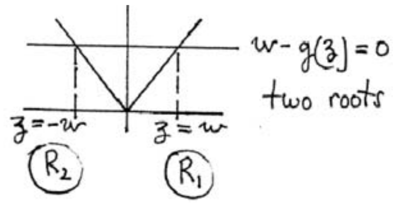


Figure 20:

Consider $w = g(z) = |z|$ (Fig. 20): At root R_1: $g'(z)|_w = 1$; At root R_2: $g'(z)|_{-w} = -1$. Then

$$f_W(w) = \begin{cases} f_Z(w)/1 + f_Z(-w)/1, & \text{for } w \geq 0; \\ 0, & \text{for } w < 0. \end{cases}$$

Hence if $f_Z(w) = f_Z(-w)$ as it is for here, then finally

$$\begin{aligned} f_w(W) &= 2f_Z(w)u(w) \\ &= \alpha e^{-\alpha w}u(w). \end{aligned}$$

26. We are told here that X and Y are continuous random variables and are independent. We are asked for the distribution function and probability density function of $Z \triangleq \min(X, Y)$. For the distribution function, we write

$$\begin{aligned} F_Z(z) &\triangleq P[Z \leq z] \\ &= 1 - P[Z > z] \\ &= 1 - P[X > z]P[Y > z] \\ &= 1 - (1 - P[X \leq z])\,(1 - P[Y \leq z]) \\ &= 1 - (1 - F_X(z)\,(1 - F_Y(z)) \\ &= F_X(z) + F_Y(z) - F_X(z)F_Y(z). \end{aligned}$$

For the pdf, we can write

$$\begin{aligned} f_Z(z) &\triangleq \frac{dF_Z(z)}{dz} \\ &= \frac{d\,(F_X(z) + F_Y(z) - F_X(z)F_Y(z))}{dz} \\ &= f_X(z) + f_Y(z) - f_X(z)F_Y(z) - F_X(z)f_Y(z). \end{aligned}$$

For X and Y uniformly distributed as $U[0,1]$, we then get,

$$\begin{aligned} F_Z(z) &= z + z - z^2 \quad \text{for } z \in [0,1], \text{ or generally,} \\ &= \begin{cases} 2z - z^2, & 0 \leq z \leq 1, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

The pdf becomes,

$$\begin{aligned} f_Z(z) &= 1 + 1 - 2z \quad \text{for } z \in [0,1], \text{ or generally,} \\ &= \begin{cases} 2 - 2z, & 0 \leq z \leq 1, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

Here are the plots:

$f_Z(z) = 2 - 2z$ (Fig. 21):

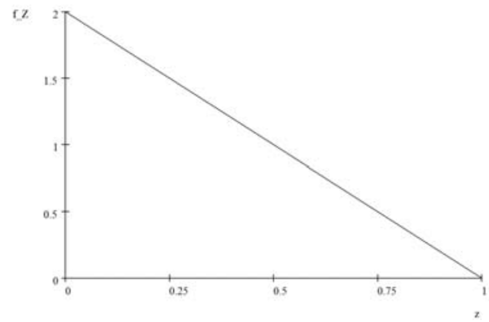


Figure 21:

$F_Z(z) = 2z - z^2$ (Fig. 22):

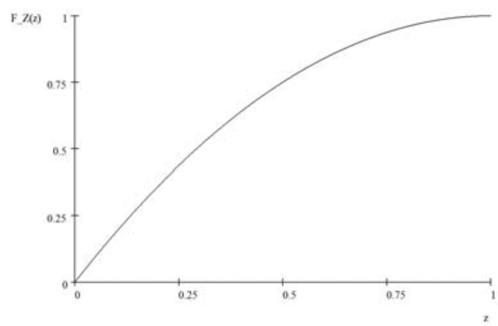


Figure 22:

Next we repeat the plots for the case where X and Y are independent and exponentially distributed with parameter $\alpha(>0)$, i.e. $f_X(x) = f_Y(x) = \alpha \exp(-\alpha x)u(x)$. Plugging into the equations, we get $F_Z(z) = (1 - \exp(-2\alpha z))u(z)$ and $f_Z(z) = 2\alpha \exp(-2\alpha z)u(z)$. We notice that, unlike the case of the two uniform random variables, the minimum of two independent exponential random variables remains exponential, but with twice the parameter value. The plots then become, for $\alpha = 2$,

$f_Z(z) = 4 \exp(-4z)$ (Fig. 23):

$F_Z(z) = 1 - \exp(-4z)$ (Fig. 24):

27.

$$\begin{aligned}
F_Z(z) &= F_{X_1,X_2}(z,z) \\
&= F_{X_1}(z)F_{X_2}(z) \\
&= [1 - \exp(-z/\mu)]\,[1 - \exp(-z/\mu)]\;\; u^2(z) \\
&= [1 - \exp(-z/\mu)]^2\;\; u(z) \\
&\ (1 - e^{-z/\mu})^2 u(z).
\end{aligned}$$

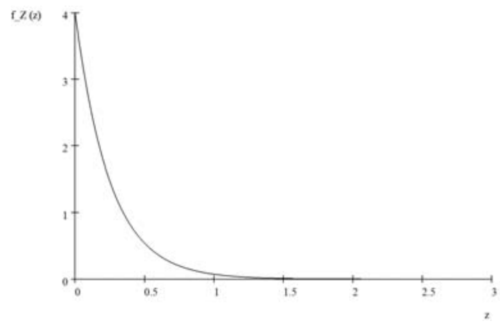


Figure 23:

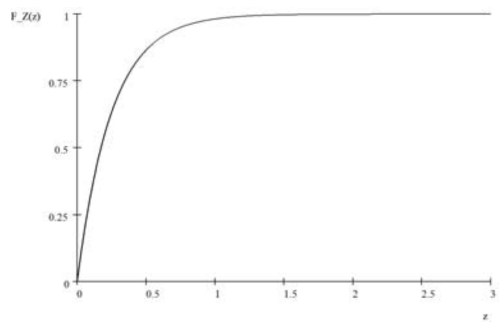


Figure 24:

$$
\begin{aligned}
f_Z(z) &= \frac{d}{dz} F_Z(z) \\
&= \frac{d}{dz}\left([1 - \exp(-z/\mu)]^2 \ u(z) \right) \\
&= \frac{d}{dz}\, [1 - \exp(-z/\mu)]^2 \ u(z) \\
&= \frac{2}{\mu} \exp(-z/\mu)\, [1 - \exp(-z/\mu)] \ u(z) \\
&= \frac{2}{\mu} e^{-z/\mu} (1 - e^{-z/\mu}) u(z)
\end{aligned}
$$

This problem has an interpretation for the time to failure of two independent machines. If we regard the two random variables X_1 and X_2 as waiting times (to failure), then Z would be the time till both machines fail. This has clear application in high-reliability computing.

28. The event $\{Z \leq z\}$ is the same as

$$
\begin{aligned}
& \{\max(X_1, X_2, ..., X_n) \leq z\} \\
= \ & \{X_1 \leq z, X_2 \leq z, ..., X_n \leq z\}
\end{aligned}
$$

Since $\{X_i\}$ are independent, then

$$F_Z(z) = F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z).$$

Since $\{X_i\}$ are i.i.d.,

$$F_{X_1}(z) = F_{X_2}(z) = \cdots = F_{X_n}(z),$$

so that

$$\begin{aligned} F_Z(z) &= (F_{X_1}(z))^n \\ &= F_{X_1}^n(z). \end{aligned}$$

29. The event $\{Z > z\} = \{\min(X_1, X_2, ..., X_n) > z\}$ is identical with

$$\{Z > z\} = \{X_1 > z, X_2 > z, ..., X_n > z\}.$$

Thus

$$\begin{aligned} P[Z > z] &= P[\{X_1 > z, X_2 > z, ..., X_n > z\}] \\ &= P[X_1 > z]P[X_2 > z]\cdots P[X_n > z] \\ &= (1 - F_{X_1}(z))(1 - F_{X_2}(z))\cdots(1 - F_{X_n}(z)). \end{aligned}$$

Because the $\{X_i\}$ are i.i.d., this becomes

$$P[Z > z] = (1 - F_{X_1}(z))^n,$$

and so for the complementary event $\{Z \leq z\}$, we have

$$F_Z(z) = 1 - (1 - F_{X_1}(z))^n.$$

30. $Z_n \triangleq \max(X_1, X_2, ..., X_n)$ and the X_is are independent RVs. Then

$$\begin{aligned} F_{Zn}(z) &= P[Z_n \leq z] \\ &= P[X_1 \leq z]P[X_2 \leq z]\cdots P[X_n \leq z] \\ &= (F_X(z))^n \\ &= (1 - e^{-z})^n u(z). \end{aligned}$$

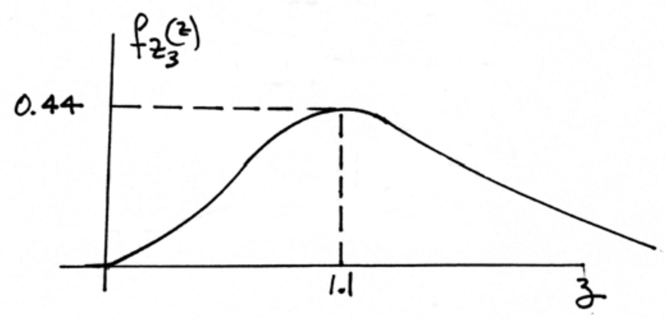
Hence

$$\begin{aligned} f_{Z_n}(z) &= \frac{dF_{Zn}(z)}{dz} \\ &= n((1 - e^{-z})^{n-1} e^{-z} u(z). \end{aligned}$$

The peak of this curve will occur at

$$\begin{aligned} 0 &= f'_{Z_n}(z) = \\ &= n(n-1)(1 - e^{-z})^{n-2} e^{-2z} - n(1 - e^{-z})^{n-1} e^{-z} \\ &= (n-1)e^{-z} - (1 - e^{-z}), \end{aligned}$$

which happens at $ne^{-z_o} = 1$ or $z_o = \ln(n)$. For $n = 3$, $z_o \simeq 1.1$ and $f_{Z_3}(z_o) \simeq 0.444$. See sketch below for $n = 3$.



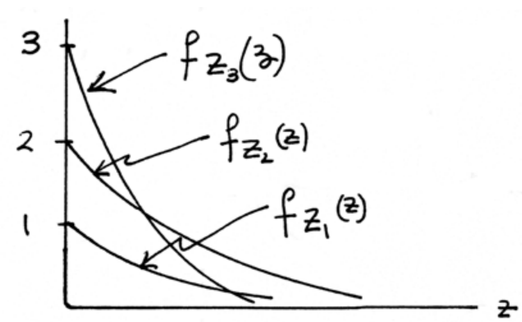
As for the minimum, we have then $Z_n \triangleq \max(X_1, X_2, ..., X_n)$, where again the X_is are independent RVs. Then

$$
\begin{aligned}
F_{Z_n}(z) &= P[Z_n \le z] \\
&= 1 - P[Z_n > z] \\
&= 1 - P[X_1 > z]P[X_2 > z] \cdots P[X_n > z] \\
&= 1 - (1 - F_{X_1}(z))(1 - F_{X_2}(z)) \cdots (1 - F_{X_n}(z)) \\
&= 1 - [1 - (1 - e^{-z})u(z)]^n \\
&= \begin{cases} 0, & z < 0, \\ (1 - e^{-nz}) & z \ge 0 \end{cases} \\
&= (1 - e^{-nz})u(z).
\end{aligned}
$$

So the pdf is given as

$$
\begin{aligned}
f_{Z_n}(z) &= F'_{Z_n}(z) \\
&= ne^{nz}u(z).
\end{aligned}
$$

Here is a sketch for $n = 1, 2, 3$.

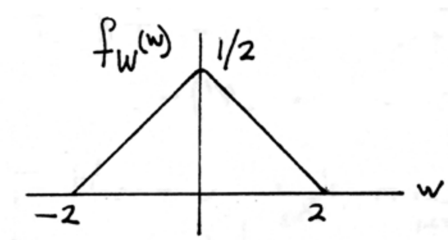


31. We are given $W = X + Y$, where X, Y are independent and identically distributed (i.i.d.)

with $X : U(-1,+1) : Y$. Then

$$\begin{aligned} f_W(w) &= f_X(w) * f_Y(w) \\ &= f_X(w) * f_X(w) \\ &= \mathrm{rect}(\frac{w}{2}) * \mathrm{rect}(\frac{w}{2}) \\ &= \left(\frac{1}{2} - \frac{|w|}{4}\right) \mathrm{rect}(\frac{w}{4}), \end{aligned}$$

with sketch:



Next, consider

$$g(w) = \begin{cases} \sqrt{w}, & w \geq 0, \\ 0, & w < 0, \end{cases}$$

with sketch:

Now, the real roots of $z - g(w) = 0$ are just one $w = z^2, z > 0$, and no real roots when $z < 0$. Hence

$$f_Z(z) = \frac{f_W(z^2)}{|g'(w)|_{w=z^2}},$$

where $g'(w) = \frac{1}{2\sqrt{w}}$ and $|g'(z^2)| = \frac{1}{2z}$. Thus, we have

$$f_Z(z) = \begin{cases} z - \frac{z^3}{2}, & 0 < z \leq \sqrt{2}, \\ 0, & \text{else}. \end{cases}$$

Finally, for $z = 0$, we have the probability mass

$$\begin{aligned} P[Z=0] &= P[W \leq 0] \\ &= \int_{-2}^{0} f_W(w) dw \\ &= \int_{-2}^{0} \left(\frac{1}{2} - \frac{|w|}{4}\right) dw \\ &= \frac{1}{2}. \end{aligned}$$

We can write this as a delta function in the pdf f_Z as

$$f_Z(z) = \frac{1}{2}\delta(z) + \left(z - \frac{z^3}{2}\right)[u(z) - u(z-\sqrt{2})].$$

32. We have from the problem statement that

$$\begin{aligned} Z &= \alpha X_1 + \alpha X_2 \\ &= X_1' + X_2', \quad \text{with } X_i' \triangleq \alpha X_i \;\; \text{for } i = 1,2. \end{aligned}$$

The pdf's of the primed variables become

$$f_{X_i'}(x') = \frac{1}{\alpha} f_{X_i}(x'/\alpha),$$

and since the X_i, and hence the X_i', are independent, we will have

$$\begin{aligned} f_Z(z) &= f_{X_1'}(z) * f_{X_2'}(z) \\ &= \frac{1}{\alpha^2} f_{X_1}(z/\alpha) * f_{X_2}(z/\alpha). \end{aligned}$$

Now we are given that the X_i are distributed as $U[0,b]$, thus $f_{X_i}(x) = \frac{1}{b}[u(x) - u(x-b)]$, thus

$$\begin{aligned} f_{X_1'}(z) &= \frac{1}{\alpha b}[u(\frac{x}{\alpha}) - u(\frac{x}{\alpha} - b)] \\ &= \frac{1}{\alpha b}[u(x) - u(x - \alpha b)], \quad \text{since } \alpha > 0, \end{aligned}$$

and so

$$f_Z(z) = \frac{1}{\alpha b}[u(z) - u(z-\alpha b)] * \frac{1}{\alpha b}[u(z) - u(z - \alpha b)].$$

Performing this convolution, we get a triangle with support $[0, 2\alpha b]$ on the z axis and height $1/\alpha b$.We can then write down the pdf of Z as

$$f_Z(z) = \begin{cases} \frac{1}{(\alpha\beta)^2} z, & 0 < z \leq \alpha b, \\ \frac{1}{\alpha b}\left(2 - \frac{1}{\alpha\beta} z\right), & \alpha b < z \leq 2\alpha b, \\ 0, & \text{else.} \end{cases}$$

In this problem, we are given $\alpha = 1/10$ and $b = 100$, so that $\alpha b = 10$, and

$$\begin{aligned} f_Z(z) &= \frac{1}{10}[u(z) - u(z-10)] * \frac{1}{10}[u(z) - u(z-10)] \\ &= \begin{cases} \frac{1}{100} z, & 0 < z \leq 10, \\ \frac{1}{10}\left(2 - \frac{1}{10} z\right), & 10 < z \leq 20, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

We can now calculate the probability that the plane will fly at least 5 hours (area of shaded region in Fig. 25), i.e.

$$\begin{aligned} P[Z &\geq 5] = 1 - F_Z(5), \quad \text{since } Z \text{ is a continuous random variable,} \\ &= 1 - \int_0^5 \frac{1}{100} z dx \\ &= 1 - \frac{1}{100}(\frac{1}{2}z^2|_0^5) \\ &= 1 - \frac{1}{8} = 0.875. \end{aligned}$$

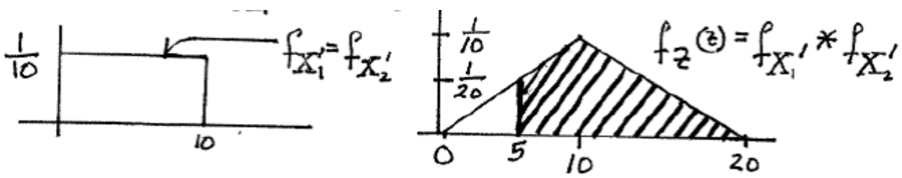


Figure 25:

33. Let $Z = X + Y$, then

$$\begin{aligned} P_Z(n) &= P[Z = n] \\ &= \sum_{k=0}^{\infty} P_X(k) P_Y(n-k) \\ &= e^{-2} e^{-3} \sum_{k=0}^{n} \frac{1}{k!} \frac{1}{(n-k)!} 2^k 3^{n-k} \\ &= \frac{e^{-5}}{n!} \sum_{k=0}^{n} \binom{n}{k} 2^k 3^{n-k} \\ &= \frac{e^{-5}}{n!} 5^n, n \geq 0. \end{aligned}$$

Hence, Z is Poisson distributed with parameter $\lambda_Z = \lambda_X + \lambda_Y$. So

$$\begin{aligned} P[Z \leq 5] &= \sum_{k=0}^{5} P_Z(k) \\ &= e^{-5} \left(1 + 5 + \frac{5^2}{2} + \frac{5^3}{6} + \frac{5^4}{24} + \frac{5^5}{120} \right) \\ &\simeq 0.616 \end{aligned}$$

34. This is an invertible transformation, with inverse

$$x = u + v \qquad \text{and} \qquad y = u - v.$$

The joint density $f_{U,V}$ is then

$$f_{U,V}(u, v) = f_{X,Y}(u + v, u - v)\ |\widetilde{J}|,$$

where the Jacobian of the inverse transformation $\widetilde{J}$ is determined as

$$\begin{aligned} \widetilde{J} &= \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \\ &= \det \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= -2. \end{aligned}$$

Thus we have

$$\begin{aligned}
f_{U,V}(u,v) &= f_{X,Y}(u+v, u-v)\, |\widetilde{J}| \\
&= 2 f_{X,Y}(u+v, u-v) \\
&= 2 f_X(u+v) f_Y(u-v), \\
&= \begin{cases} 2, & 0 < u+v < 1 \text{ and } 0 < u-v < 1, \\ 0, & \text{else,} \end{cases}
\end{aligned}$$

since X and Y are independent and uniformly distributed $U(0,1)$. We can also write the answer in terms of the *unit pulse function* $\mathrm{rect}(\cdot)$ defined as

$$\mathrm{rect}(x) \triangleq \begin{cases} 1, & -0.5 < x < +0.5, \\ 0, & \text{else.} \end{cases}$$

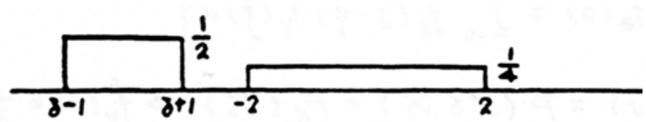
The result can then be written as

$$f_{U,V}(u,v) = 2\,\mathrm{rect}(u+v-0.5)\,\mathrm{rect}(u-v-0.5).$$

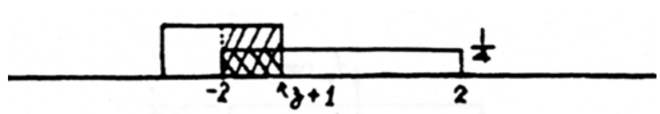
35. (a) $Z = X + Y$ where X and Y are independent, and $X : U(-1,+1)$, i.e. $f_X(x) = \frac{1}{2}\mathrm{rect}(\frac{x}{2})$ and $Y : U(-2,+2)$, i.e. $f_X(x) = \frac{1}{4}\mathrm{rect}(\frac{x}{4})$. Then

$$\begin{aligned}
f_Z(z) &= f_X(z) * f_Y(z) \\
&= \int_{-\infty}^{\infty} f_X(z-v) f_Y(v)\, dv \\
&= \int_{-\infty}^{\infty} \frac{1}{2}\mathrm{rect}(\frac{z-v}{2}) \frac{1}{4}\mathrm{rect}(\frac{v}{4})\, dv \\
&= \frac{1}{8}\int_{-2}^{+2} \mathrm{rect}(\frac{z-v}{2})\, dv.
\end{aligned}$$

To perform this convolution, we start in *region 1*, were $z < -3$ and there is no overlap between the two function supports. Here is the sketch.



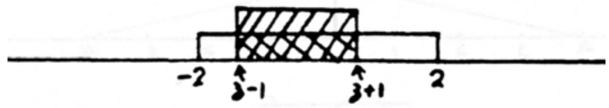
Clearly there is no overlap here in region 1, and so the output $f_Z(z) = 0$ here.
region 2: partial overlap $-3 \le z < -1$. Here is the sketch:



Using the cross-hatched area as a guide, we can evaluate the integral as

$$\begin{aligned}
f_Z(z) &= \frac{1}{8}\int_{-2}^{z+1} dv \\
&= \frac{1}{8}(z+3), \quad -3 \le z < -1.
\end{aligned}$$

Next comes region 3, complete overlap, sketched as



region 3: $-1 \leq z < +1$. From the sketch, we can see that the integral must start at $z-1$ and go to $z+1$ in this region. Thus

$$\begin{aligned} f_Z(z) &= \frac{1}{8}\int_{z-1}^{z+1} dv \\ &= \frac{1}{4}. \end{aligned}$$

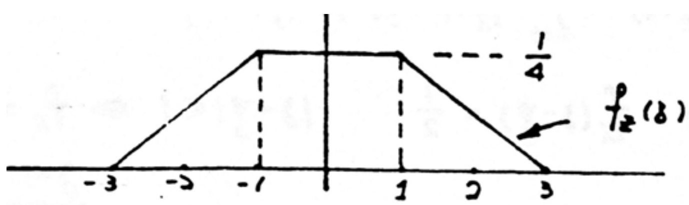
Next comes region 4, partial overlap on the right, where $z-1 < 2$ and $z+1 \geq 2$, which together imply $1 \leq z < 3$.
region 4: $1 \leq z < 3$. Here

$$\begin{aligned} f_Z(z) &= \frac{1}{8}\int_{z-1}^{2} dv \\ &= \frac{1}{8}(3-z), \quad 1 \leq z < 3. \end{aligned}$$

Finally comes region 5, no overlap on right.
region 5: $z \geq 3 \qquad f_Z(z) = 0$.
Our complete answer can be written as

$$f_Z(z) = \begin{cases} 0, & z < -3 \\ \frac{1}{8}(z+3), & -3 \leq z < -1, \\ \frac{1}{4} & -1 \leq z < +1, \\ \frac{1}{8}(3-z), & 1 \leq z < 3, \\ 0, & z \leq 3, \end{cases}$$

with sketch:

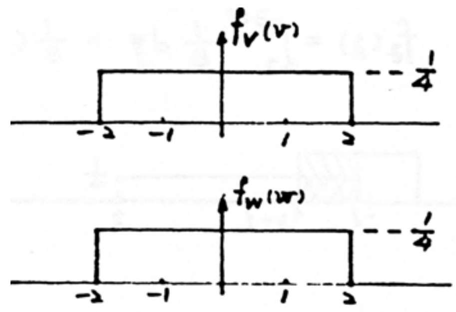


As a check, we note that the total area is $2(\frac{1}{2}(2)\frac{1}{4}) + 2(\frac{1}{4}) = \frac{1}{2} + \frac{1}{2} = 1$.

(b) Here $Z = 2X - Y$, where again X and Y are independent. We let $V = 2X$ and $W = -Y$ to make use of convolution of densities. Now $F_V(v) = P[V \leq v] = P[X \leq \frac{v}{2}] = F_X(\frac{v}{2}) \Longrightarrow f_V(v) = \frac{1}{2}f_X(\frac{v}{2})$ and $F_W(w) = P[W \leq w] = P[Y > -w] = 1 - F_Y(-w) \Longrightarrow f_W(w) = f_Y(-w)$. Hence

$$f_V(v) = \frac{1}{4}\text{rect}(\frac{v}{4}) \quad \text{and} \quad f_W(w) = \frac{1}{4}\text{rect}(\frac{w}{4}),$$

with sketches



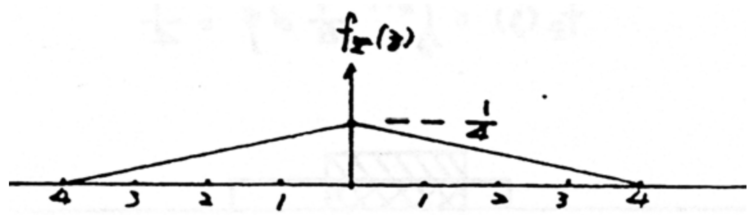
so

$$
\begin{aligned}
f_Z(z) &= f_V(z) * f_W(z) \\
&= \frac{1}{8}\mathrm{rect}(\frac{z}{4}) * \mathrm{rect}(\frac{z}{4}) \\
&= \frac{1}{4}\mathrm{triag}(\frac{z}{4}),
\end{aligned}
$$

where function 'triag' is given as

$$
\mathrm{triag}(x) \triangleq \begin{array}{ll} (1-|x|), & |x| \leq 1, \\ 0, & |x| > 1. \end{array}
$$

A sketch of this f_Z is given as



36. We look at the transformation problem for two independent Normal random variables X and $Y : N(0, \sigma^2)$, transformed to $Z \triangleq X^2 + Y^2$ and $W \triangleq X$. We thus have

$$
z = g(x, y) = x^2 + y^2 \quad \text{and} \quad w = h(x, y) = x.
$$

This is a non-invertible transformation with two real roots, for $|w| < \sqrt{z}, z > 0$,

$$
\begin{aligned}
R_1 &: \quad x = w, y = +\sqrt{z - w^2}, \text{ and} \\
R_2 &: \quad x = w, y = -\sqrt{z - w^2}.
\end{aligned}
$$

Now at both roots the magnitude of the Jacobian is the same,

$$
\begin{aligned}
|J_1| &= |J_2| = 2\sqrt{z - w^2}, \text{ where} \\
J_{1,2} &= \det \begin{bmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{bmatrix} = 2y = \pm 2\sqrt{z - w^2}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
f_{Z,W}(z,w) &= \frac{1}{2\sqrt{z-w^2}}\left(f_{X,Y}(w,\sqrt{z-w^2}) + f_{X,Y}(w,-\sqrt{z-w^2})\right) \\
&= \begin{cases} \frac{1}{2\pi\sigma^2}\frac{1}{\sqrt{z-w^2}}\exp(-z/2\sigma^2), & |w| < \sqrt{z}, z > 0 \\ 0, & \text{else.} \end{cases}
\end{aligned}
$$

We can find the marginal density f_Z either by integrating out the unwanted variable in this joint density, or by using the result of Example 3.3-10 that Z will be Exponential distributed (equivalently Chi-square with 2 degrees of freedom). Either way the answer is

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{+\infty} f_{Z,W}(z,w)dw \\
&= \frac{1}{2\pi\sigma^2}\int_{-\sqrt{z}}^{+\sqrt{z}} \frac{1}{\sqrt{z-w^2}}\exp(-z/2\sigma^2)dw \\
&= \frac{1}{2\sigma^2}\exp(-z/2\sigma^2)\, u(z).
\end{aligned}
$$

37. We want to find the pdf of the two variables

$$Z = aX + bY$$

$$W = cX + dY$$

The joint pdf of X and Y are given as

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}e^{-Q(x,y)},$$

where $Q(x,y) = \frac{1}{2\sigma^2(1-\rho^2)}[x^2 - 2\rho xy + y^2]$. Consider the inverse transformation, i.e.,

$$X = \hat{a}Z + \hat{b}W$$

$$Y = \hat{c}Z + \hat{d}W$$

Note that from the above transformation, the solution of the two equations are given as $x = \hat{a}z + \hat{b}w$ and $y = \hat{c}z + \hat{d}w$. The term in the exponent of the pdf will be $\frac{1}{2\sigma^2(1-\rho^2)}[x^2 - 2\rho xy + y^2]$, and this is given as

$$\frac{1}{2\sigma^2(1-\rho^2)}[x^2 - 2\rho xy + y^2] = \frac{1}{2\sigma^2(1-\rho^2)}\left[(\hat{a}z + \hat{b}w)^2 - 2\rho(\hat{a}z + \hat{b}w)(\hat{c}z + \hat{d}w) + (\hat{c}z + \hat{d}w)^2\right].$$

In this exponent, if the cross terms (terms that contain zw) vanish, then we would be able to split the pdf in to the product of the two marginal pdf's. In other words, if the coefficients of the terms zw is zero, then we would be able to write $f_{Z,W} = f_Z f_W$. Therefore, we need

$$(2\hat{a}\hat{b} + 2\hat{c}\hat{d} - 2\rho\hat{a}\hat{d} - 2\rho\hat{b}\hat{c}) = 0$$

. If we chose $\hat{a} = \hat{c}, \hat{b} = -\hat{d}$, the coefficient of zw will be zero. Then

$$x = \hat{a}z + \hat{b}w$$

$$y = \hat{a}z - \hat{b}w$$

will give us $\frac{1}{2\sigma^2(1-\rho^2)}[x^2 + y^2 - 2\rho xy] = \frac{1}{2\sigma^2(1-\rho^2)}[2(1-\rho)\hat{a}^2z^2 + 2(1+\rho)\hat{b}^2w^2]$. Therefore, $Q(x,y) = \frac{1}{2}\left\{\left[\frac{z\sqrt{1-\rho}}{\sigma/\sqrt{2}\hat{a}}\right]^2 + \left[\frac{w\sqrt{1+\rho}}{\sigma/\sqrt{2}\hat{b}}\right]^2\right\}$. The magnitude of the Jacobian is given as

$$|J| = \text{mag}\begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} = \frac{1}{2|\hat{a}\hat{b}|},$$

where $g(x,y) = \frac{x+y}{2\hat{a}}$, $h(x,y) = \frac{x-y}{2\hat{b}}$. With $\sigma_1 \triangleq \sigma\sqrt{1-\rho}/\sqrt{2}\hat{a}$, $\sigma_2 \triangleq \sigma\sqrt{1+\rho}/\sqrt{2}\hat{b}$. Hence,

$$\begin{aligned} f_{ZW}(z,w) &= \frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\frac{z^2}{\sigma_1^2}}\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\frac{w^2}{\sigma_2^2}} \\ &= f_Z(z)f_W(w). \end{aligned}$$

38. Let $g(x,y) \triangleq \frac{x^2+y^2}{2}$ and $h(x,y) \triangleq \frac{x^2-y^2}{2}$. The real roots of $g(x,y) = v$, $h(x,y) = w$ occur for $v \geq 0, |v| \geq |w|$ and are four in number.

$$\begin{aligned} x_1 &= +\sqrt{v+w}, & y_1 &= +\sqrt{v-w} \\ x_2 &= -\sqrt{v+w}, & y_2 &= +\sqrt{v-w} \\ x_3 &= -\sqrt{v+w}, & y_3 &= -\sqrt{v-w} \\ x_4 &= +\sqrt{v+w}, & y_4 &= -\sqrt{v-w}. \end{aligned}$$

We note that w can be negative, but never greater in magnitude than v. The magnitude of J is

$$\text{abs}\begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} = 2|xy| = 2\sqrt{v^2 - w^2},$$

and this is the same for all the four roots. Now observe: $x^2 + y^2 = 2v$ for all the roots, and

$$x_1y_1 = \sqrt{v^2 - w^2} = x_3y_3$$

$$x_2y_2 = -\sqrt{v^2 - w^2} = x_4y_4$$

Hence

$$\begin{aligned} f_{VW}(v,w) &= \frac{1}{4\pi\sqrt{v^2-w^2}\sqrt{1-\rho^2}}2\left\{e^{-\left[\frac{2v-2\sqrt{v^2-w^2}\rho}{2(1-\rho^2)}\right]} + e^{-\left[\frac{2v+2\sqrt{v^2-w^2}\rho}{2(1-\rho^2)}\right]}\right\} \\ &= \frac{1}{\pi\sqrt{(1-\rho^2)(v^2-w^2)}}e^{-\frac{v}{1-\rho^2}}\cosh\left(\frac{\rho\sqrt{v^2-w^2}}{1-\rho^2}\right), \quad \text{for } v \geq 0, |v| \geq |w|, \end{aligned}$$

where the hyperbolic cosine function $\cosh(x) \triangleq \frac{1}{2}(e^x + e^{-x})$. For $v < 0$ or $|v| < |w|$, there are no real roots of the transformation equations, so that we have $f_{VW}(v,w) = 0$ there.

39. From Eq. 3.4-4 in Example 3.4-1, we have the linear transformation

$$v = x + y, w = x - y,$$

with the one solution, i.e. $n = 1$:

$$x = \phi(v, w) = \frac{v + w}{2}, y = \psi(v, w) = \frac{v - w}{2}.$$

Proceeding with the direct approach, the matrix in the Jacobian $\widetilde{J}$ becomes:

$$\begin{bmatrix} \frac{\partial \phi}{\partial v} & \frac{\partial \phi}{\partial w} \\ \frac{\partial \psi}{\partial v} & \frac{\partial \psi}{\partial w} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix},$$

with determinant or Jacobian given as

$$\widetilde{J} = \left| \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \right| = -\frac{1}{2}.$$

So, the magnitude of the Jacobian is $\frac{1}{2}$. Then by Eq. 3.4-11, we can write

$$\begin{aligned} f_{VW}(v, w) &= |\widetilde{J}| f_{XY}\left(\frac{v + w}{2}, \frac{v - w}{2}\right) \\ &= \frac{1}{2} f_{XY}\left(\frac{v + w}{2}, \frac{v - w}{2}\right), \end{aligned}$$

which agrees with Eq. 3.4-4 that was derived in the Example 3.4-1 using the indirect method.

40. Here the transformation is give as

$$\begin{aligned} z &= g(x, y) \triangleq x \cos\theta + y \sin\theta, \qquad (1) \\ w &= h(x, y) \triangleq x \sin\theta - y \cos\theta. \qquad (2) \end{aligned}$$

The random variables X and Y are independent Normal distributed as $N(0, 1)$. To find the inverse, we can multiply (1) by $\cos\theta$ and then multiply (2) by $\sin\theta$ and add the results to get

$$x = z \cos\theta + w \sin\theta,$$

since $\cos^2\theta + \sin^2\theta = 1$. Similarly, we multiply (1) by $\sin\theta$ and multiply (2) by $\cos\theta$ and subtract to get

$$y = z \sin\theta - w \cos\theta.$$

Thus there is a unique inverse and one root at $(x, y) = (z \cos\theta + w \sin\theta, z \sin\theta - w \cos\theta)$. Calculating the Jacobian, we get

$$J = \det \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} = -1, \quad \text{so that } |J| = 1.$$

The transformed pdf then becomes,

$$\begin{aligned} f_{Z,W}(z, w) &= f_{X,Y}(z \cos\theta + w \sin\theta, z \sin\theta - w \cos\theta) \cdot 1 \\ &= \frac{1}{2\pi} \exp -\frac{1}{2} \left((z \cos\theta + w \sin\theta)^2 + (z \sin\theta - w \cos\theta)^2 \right) \\ &= \frac{1}{2\pi} \exp -\frac{1}{2} \left(z^2 + w^2 \right), \qquad -\infty < z, w < +\infty. \end{aligned}$$

This independent Normal joint density is thus unchanged by this transformation! Note that this transformation is closely related to a rotation of Cartesian coordinates in the $x-y$ plane. In fact, if w were replaced in the transformation by $-w$, the transformation would be a coordinate rotation by angle $+\theta$ in the $x-y$ plane.

41. We look at the transformation problem for two independent Normal random variables X and $Y\ : N(0,\sigma^2)$, transformed to $Z \triangleq X^2+Y^2$ and $W \triangleq 2Y$. We thus have

$$z = g(x,y) = x^2+y^2 \quad \text{and} \quad w = h(x,y) = 2y.$$

This is a non-invertible transformation with two real roots, for $|w| < \sqrt{z}, z>0$,

$$\begin{aligned} r_1 &: \quad x = +\sqrt{z-\left(\frac{w}{2}\right)^2}, y = \frac{w}{2}, \text{ and} \\ r_2 &: \quad x = -\sqrt{z-\left(\frac{w}{2}\right)^2}, y = \frac{w}{2}. \end{aligned}$$

Now at both roots the magnitude of the Jacobian is the same,

$$\begin{aligned} |J_1| &= |J_2| = 4\sqrt{z-\left(\frac{w}{2}\right)^2}, \text{ where} \\ J_{1,2} &= \det\begin{bmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{bmatrix} = 4x = \pm 4\sqrt{z-\left(\frac{w}{2}\right)^2}. \end{aligned}$$

Hence

$$\begin{aligned} f_{Z,W}(z,w) &= \frac{1}{4\sqrt{z-\left(\frac{w}{2}\right)^2}}\left(f_{X,Y}(\sqrt{z-\left(\frac{w}{2}\right)^2},\frac{w}{2}) + f_{X,Y}(-\sqrt{z-\left(\frac{w}{2}\right)^2},\frac{w}{2})\right) \\ &= \begin{cases} \frac{1}{4\pi\sigma^2}\frac{1}{\sqrt{z-\left(\frac{w}{2}\right)^2}}\exp(-z/2\sigma^2), & |\frac{w}{2}| < \sqrt{z}, z>0 \\ 0, & \text{else.} \end{cases} \end{aligned}$$

42. First we determine the value of A

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x,y)dxdy \\ &= A\int_0^1\int_{-1}^{+1}(x^2+y^2)dxdy \\ &= A\int_0^1\left(\int_{-1}^{+1}(x^2+y^2)dy\right)dx \\ &\Rightarrow \ A = \frac{3}{4} \end{aligned}$$

Hence $f_{XY}(x,y) = \frac{3}{4}(x^2+y^2)$ with support on $0 \le x \le 1, -1 \le y \le +1$.

(a) i) For $x<0, y<-1$,

$$F_{XY}(x,y) = 0,$$

i) For $x > 1, y > 1$,

$$F_{XY}(x,y) = 1.$$

iii) $\quad 0 \le x \le 1, y > 1$,

$$\begin{aligned} F_{XY}(x,y) &= \frac{3}{4}\int_0^x \left(\int_{-1}^{+1} (u^2+v^2)dv\right) du \\ &= \frac{3}{4}\left[\frac{x^3}{3}y + \frac{y^3}{3}x\Bigg|_{-1}^{+1}\right] \\ &= \frac{3}{4}\left[\frac{x^3}{3} + \frac{x}{3} - \left(-\frac{x^3}{3} - \frac{x}{3}\right)\right] \\ &= \frac{3}{4}\left(\frac{2}{3}x^3 + \frac{2}{3}x\right) \\ &= \frac{1}{2}(x^3+x). \end{aligned}$$

iv) For $0 \le x \le 1, -1 < y < 1$,

$$\begin{aligned} F_{XY}(x,y) &= \frac{3}{4}\int_0^x \left(\int_{-1}^{y} (u^2+v^2)dv\right) du \\ &= \frac{1}{4}\left[x^3(y+1) + x(y^3+1)\right] \end{aligned}$$

v) For $x > 1, -1 < y < 1$,

$$\begin{aligned} F_{XY}(x,y) &= \frac{3}{4}\int_0^1 \left(\int_{-1}^{y} (u^2+v^2)dv\right) du \\ &= \frac{1}{4}\left[(y+1) + (y^3+1)\right]. \end{aligned}$$

43. (a) The range of Y, R_Y is a subset of the real line R^1 if Y is a real-valued random variable.

(b) A reasonable probability space for X is $(\Omega, \mathcal{F}, P)$ where Ω is the sample space of the underlying experiment, $\mathcal{F}$ is the $\sigma-$field of events defined on Ω, and P is the set function that assigns to every set (event) $E \in \mathcal{F}$, the number $P[E]$. A reasonable probability space for Y is $(R_X, \mathcal{B}, P_X)$ where R_X, the range of X, is the induced sample space under the mapping X, $\mathcal{B}$ is the Borel field of events over R^1, and P_X is a set function that assigns to every set $B \in \mathcal{B}$, the number $P_X[B]$.

(c) The event $\{\varsigma | Y(\varsigma) \le y\}$ under the mapping Y is the set $(-\infty, y] \in R^1$.

(d) In the I/O viewpoint, the inverse image is computed as follows:

$$\begin{aligned} \{Y \le y\} &= \{2X + 3 \le y\} \\ &= \left\{X \le \frac{y-3}{2}\right\} \\ &= \left(-\infty, \frac{y-3}{2}\right]. \end{aligned}$$

44. Define the event $A = \{\max(T_1, T_2) \leq t\}$. Then

$$F_Y(y,t) = P[Y \leq y|A]P[A] + P[Y \leq y|A^c]P[A^c].$$

Now

$$\begin{aligned} P[A] &= P[T_1 \leq t]P[T_2 \leq t], \quad \text{by independence,} \\ &= (1 - e^{-\lambda t})u(t)(1 - e^{-\lambda t})u(t) \\ &= (1 - e^{-\lambda t})^2 u(t), \end{aligned}$$

and so $P[A^c] = 1 - (1 - e^{-\lambda t})^2 u(t)$. Observe that $Y = 0$ at t given event A, while $Y = X$ at t given event A^c. So

$$P[Y \leq y|A] = u(y) \quad \text{and} \quad P[Y \leq y|A^c] = F_X(y).$$

Hence,

$$\begin{aligned} F_Y(y,t) &= P[Y \leq y|A]P[A] + P[Y \leq y|A^c]P[A^c] \\ &= u(y)(1 - e^{-\lambda t})^2 u(t) + F_X(y)(1 - (1 - e^{-\lambda t})^2 u(t)) \\ &= (1 - e^{-\lambda t})^2 u(y)u(t) + F_X(y)(1 - (1 - e^{-\lambda t})^2 u(t)). \end{aligned}$$

For any fixed $t > 0$ and $y = \infty$, we see $F_Y(\infty, t) = (1 - e^{-\lambda t})^2 + 1(1 - (1 - e^{-\lambda t})^2) = 1$ and for $t = \infty$, and any fixed y, we have $F_Y(y, \infty) = u(y)$. Does this make sense? Think about it!

Solutions to Chapter 4

1. The sample mean for the set of numbers is given by
$$X_s = \frac{1}{12}\sum_{i=1}^{12} X_i \approx 1.07.$$
The standard deviation is given by
$$\sigma_s = \sqrt{\frac{1}{12}\sum_{i=1}^{12}(X_i - X_s)^2} \approx 3.98.$$

2. We are given that X is Bernoulli distributed over $\{0,1\}$ with parameter $p(1 \geq p \geq 0)$ with the PMF
$$P_X(k) = \begin{cases} p, & k = 1, \\ 1-p, & k = 0 \\ 0, & \text{else.} \end{cases}.$$
So
$$E[X] = \sum_{k=-\infty}^{+\infty} kP_X(k) = 1p + 0(1-p) = p.$$

3. Since the random variable X is the constant value c, then its pdf $f_X(x) = \delta(x-c)$, so
$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_{-\infty}^{\infty} x\delta(x-c)dx \\ &= c. \end{aligned}$$

4. Here the random variable X is binomial distributed with parameters n a positive integer and $p(1 \geq p \geq 0)$. We are asked for the expected value $E[X]$. The PMF is given as
$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \left(u(k) - u(n+1-k)\right).$$
Thus
$$\begin{aligned} E[X] &= \sum_{k=-\infty}^{+\infty} kP_X(k) \\ &= \sum_{k=-\infty}^{+\infty} k\binom{n}{k} p^k (1-p)^{n-k} \left(u(k) - u(n+1-k)\right) \\ &= \sum_{k=0}^{n} k\binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$
To evaluate this sum, we attempt to modify it so that it looks like the binomial sum that we know. To this end, we pull out n from $\binom{n}{k}$ and pull out the variable p also, to obtain

$$
\begin{aligned}
E[X] &= \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \qquad \text{since the } k=0 \text{ term will be } 0,\\
&= np\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!}p^{(k-1)}(1-p)^{n-k}\\
&= np\sum_{k=0}^{n} \frac{(n-1)!}{(k-1)!(n-k)!}p^{(k-1)}(1-p)^{(n-1)-(k-1)}\\
&= np\sum_{i=0}^{n-1} \frac{(n-1)!}{i!(n-1-i)!}p^{i}(1-p)^{(n-1)-i} \quad \text{with the substitution } i \triangleq k-1,
\end{aligned}
$$

$$
\begin{aligned}
&= np\sum_{i=0}^{n-1} \binom{n-1}{i} p^{i}(1-p)^{(n-1)-i}\\
&= np \times (p + (1-p))^{n-1},\\
&= np \times 1 = np.
\end{aligned}
$$

5. The pdf of the uniform random variable X is given as

$$f_X(x) = \frac{1}{b-a}[u(x-a) - u(x-b)].$$

The expectation of X, which is nothing but the mean of X, is

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} x f_X(x)\,dx = \frac{1}{b-a}\int_a^b x\,dx\\
&= \frac{1}{b-a}\frac{x^2}{2}|_a^b = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2}.
\end{aligned}
$$

6. (i)

$$F_X(x) = \int_{-\infty}^{x} f_X(v)\,dv,$$

so for $x < 0$, we get $F_X(x) = 0$ since $f_X = 0$ there. For $0 \le x < 1$, we then have

$$
\begin{aligned}
F_X(x) &= \int_0^x 2v\,dv\\
&= x^2.
\end{aligned}
$$

For $x > 1$, we get $F_X(x) = F_X(1) = 1$ since $f_X = 0$ for $x > 1$. Combining these results, we have

$$F_X(x) = \begin{cases} 0, & x < 0,\\ x^2, & 0 \le x < 1,\\ 1, & x \ge 1. \end{cases}$$

(ii)

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{+\infty} x f_X(x)\, dx \\
&= \int_0^1 x 2x\, dx \\
&= 2\frac{x^3}{3}\Big|_0^1 \\
&= \frac{2}{3}.
\end{aligned}
$$

(iii)

$$
\begin{aligned}
\sigma_X^2 &= E[X^2] - (E[X])^2 \\
&= \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx - \left(\frac{2}{3}\right)^2 \\
&= \int_0^1 x^2 2x\, dx - \frac{4}{9} \\
&= 2\frac{x^4}{4}\Big|_0^1 - \frac{4}{9} \\
&= \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18}.
\end{aligned}
$$

7. The hypergeometric PMF gives the probability of x successes in k draws from a population of size $m+n$ where m is the number of success states in the population of size $m+n$. It is given by

$$
P_X(x) = \binom{m+n}{k}^{-1} \binom{m}{x} \binom{n}{k-x}
$$

and

$$
\sum_{x=0}^{k} P_X(x) = \sum_{x=0}^{k} \binom{m+n}{k}^{-1} \binom{m}{x} \binom{n}{k-x} = 1.
$$

Now consider

$$
\begin{aligned}
E[X] &= \sum_{x=0}^{k} x P_X(x) = \sum_{x=0}^{k} x \binom{m+n}{k}^{-1} \binom{m}{x} \binom{n}{k-x} \\
&= m \binom{m+n}{k}^{-1} \sum_{x=1}^{k} \binom{m-1}{x-1} \binom{n}{k-1-(x-1)}.
\end{aligned}
$$

Make the following substitutions:

$$
j \triangleq x-1, m' \triangleq m-1, \text{ and } k' \triangleq k-1,
$$

and obtain

$$
E[X] = m \binom{m+n}{k}^{-1} \sum_{j=0}^{k'} \binom{m'}{j} \binom{n}{k'-j}.
$$

The next-to-last step is to recognize that $\begin{pmatrix} m+n \\ k \end{pmatrix} = \frac{m+n}{k} \begin{pmatrix} m' \\ k' \end{pmatrix}$ so that finally

$$E[X] = m \left(\frac{m+n}{k}\right)^{-1} \begin{pmatrix} m' \\ k' \end{pmatrix}^{-1} \sum_{j=0}^{k'} \begin{pmatrix} m' \\ j \end{pmatrix} \begin{pmatrix} n \\ k'-j \end{pmatrix}.$$

But the term to the right of the product sign is unity; hence $E[X] = mk/(m+n)$.

8. By using Eq. 4.1-9, we get

$$\begin{aligned} E[g(X)] &= \int_{-\infty}^{+\infty} g(x) f_X(x) dx \\ &= \frac{1}{b-a} \int_a^b x^2 dx \\ &= \frac{1}{3(b-a)} (b^3 - a^3). \end{aligned}$$

To use Eq. 5.4-1, we must compute the density f_Y . We proceed in an indirect fashion by finding the CDF F_Y first.,

We have

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= P[X^2 \leq y] \\ &= P[-\sqrt{y} \leq X \leq \sqrt{y}] \\ &= F_X(\sqrt{y}) - F_Y(\sqrt{y}). \end{aligned}$$

Taking derivatives, we then get

$$f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}).$$

But $f_X(x) = 0$ for $x < a$ or $x > b$, hence with $a > 0$, the second term in the above equation is zero, and so

$$\begin{aligned} f_Y(y) &= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) \\ &= \frac{1}{2\sqrt{y}} \frac{1}{b-a} [u(y-a^2) - u(y-b^2)], \end{aligned}$$

where u is the unit step function. We then get for the expectation

$$\begin{aligned} E[Y] &= \int_{-\infty}^{+\infty} y f_Y(y) dy \\ &= \frac{1}{2(b-a)} \int_{a^2}^{b^2} \sqrt{y} dy \\ &= \frac{1}{2(b-a)} \left(\frac{2}{3} y^{\frac{3}{2}} \Big|_{a^2}^{b^2} \right) \\ &= \frac{1}{2(b-a)} \frac{2}{3} (b^3 - a^3) \\ &= \frac{1}{3(b-a)} (b^3 - a^3), \quad \text{the same as above.} \end{aligned}$$

9. Now $E[Y] = E[X^2+1] = E[X^2]+1$. But from Problem 4.6 we have that $E[X^2] = \sigma_X^2 + \mu_X^2 = 1/18 + 4/9 = 9/18$. Hence $E[Y] = E[X^2] + 1 = 27/18 = 1.5$. To compute σ_Y^2 we proceed as follows:

$$\begin{aligned} E[Y^2] &= E[(X^2+1)^2] \\ &= E[X^4 + 2X^2 + 1] \\ &= E[X^4] + 2 \times 9/18 + 1 \\ &= E[X^4] + 2. \end{aligned}$$

But $E[X^4] = \int_0^1 x^4(2x)dx = 1/3$. Hence $\sigma_Y^2 = E[Y^2] - E^2[Y] = 14/6 - 9/4 = 1/12$.

10. We are given $Y = X^2 + b$, therefore

$$\begin{aligned} E[Y] &= \sum_{k=0}^{\infty}(k^2+b)\frac{e^{-\alpha}\alpha^k}{k!} \\ &= e^{-\alpha}\sum_{k=0}^{\infty}k^2\frac{\alpha^k}{k!} + be^{-\alpha}\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}. \end{aligned}$$

Looking at sum in the second term, we see that

$$e^{-\alpha}\sum_{k=0}^{\infty}\frac{\alpha^k}{k!} = 1.$$

Next considering the sum in the first term above, we have

$$\begin{aligned} \sum_{k=0}^{\infty}k^2\frac{\alpha^k}{k!} &= \sum_{k=1}^{\infty}k\frac{\alpha^k}{(k-1)!} \\ &= \sum_{k=1}^{\infty}[(k-1)+1]\frac{\alpha^k}{(k-1)!} \\ &= \sum_{k=2}^{\infty}\frac{\alpha^k}{(k-2)!} + \sum_{k=1}^{\infty}\frac{\alpha^k}{(k-1)!} \\ &= \alpha^2\sum_{k=2}^{\infty}\frac{\alpha^{k-2}}{(k-2)!} + \alpha\sum_{k=1}^{\infty}\frac{\alpha^{k-1}}{(k-1)!} \\ &= (\alpha^2+\alpha)e^{\alpha}. \end{aligned}$$

Combining these results, we get

$$\begin{aligned} E[Y] &= e^{-\alpha}[\alpha^2+\alpha]e^{\alpha} + b \\ &= \alpha^2 + \alpha + b. \end{aligned}$$

11. We start from the definition

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma y + \mu) e^{-\frac{1}{2}y^2} dy, \quad \text{with } y \triangleq \frac{x-\mu}{\sigma}, \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y e^{-\frac{1}{2}y^2} dy + \mu \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}y^2} dy, \\ &= \frac{\sigma}{\sqrt{2\pi}} \times 0 + \mu \times 1 \\ &= \mu, \end{aligned}$$

where the first integral is 0 because it is an integral of an odd function over even limits, and the second integral is recognized as that of the standard Gaussian density over all its domain, and hence is 1.

12. (a) The CDF of the momentum $P = MV$ is given as

$$F_P(p) = \int_0^\infty f_V(v) \left(\int_0^{p/v} f_M(m) dm \right) dv.$$

Then we find the pdf $f_P(p)$ by differentiation as

$$\begin{aligned} f_P(p) &= dF_P(p)/dp \\ &= \int_0^\infty f_V(v) \left(f_M(m)|_{m=p/v} \right) \frac{d(p/v)}{dp} dv \\ &= \int_0^\infty f_V(v) f_M(p/v) \frac{1}{v} dv \\ &= \int_0^\infty f_M(m) f_V(p/m) \frac{1}{m} dm. \end{aligned}$$

(b) By independence $\mu_P = E[P] = E[MV] = E[M]E[V] = \mu_M \mu_V$.

13. We are to prove the general inequality $|E[X]| \leq E|X|$. We start with our definition (for continuous random variables) $E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx$. Thus

$$\begin{aligned} |E[X]| &= |\int_{-\infty}^{+\infty} x f_X(x) dx| \\ &\leq \int_{-\infty}^{+\infty} |x f_X(x)| dx \\ &= \int_{-\infty}^{+\infty} |x| f_X(x) dx \quad \text{since } f_X \geq 0, \\ &= E[|X|], \end{aligned}$$

as was to be shown.

14. If $E[g_i(X)]$ exists, then

$$\begin{aligned}
E\left[\sum_{i=1}^{n} g_i(X)\right] &= \int_{-\infty}^{+\infty} \sum_{i=1}^{n} g_i(x) f_X(x) dx \\
&= \sum_{i=1}^{n} \int_{-\infty}^{+\infty} g_i(x) f_X(x) dx \\
&= \sum_{i=1}^{n} E[g_i(X)].
\end{aligned}$$

15. This is a simple statistics calculation of mean and conditional mean.

   (a) The average number of children per household $\overline{X_S}$ is given as

$$\begin{aligned}
\overline{X_S} &= \frac{1}{20} \sum X_i = \frac{32}{20} \\
&\doteq 1.6.
\end{aligned}$$

   (b) Now, the number of households with children are 13, thus $\overline{X_{S|C}}$ (i.e. given that children are in the household) is

$$\begin{aligned}
\overline{X_{S|C}} &= \frac{32}{13} \\
&\doteq 2.46.
\end{aligned}$$

16. We know from our prior work that

$$f_X(x|B) = \begin{cases} 0, & x \leq a, \\ \frac{f_X(x)}{P[B]}, & a < x \leq b, \\ 0, & x > b. \end{cases}$$

Hence

$$E[X|B] = \int_a^b x f(x) dx / P[B].$$

Here

$$\begin{aligned}
P[B] &= F_X(b) - F_X(a) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-1}^{2} \exp(-\frac{1}{2}x^2) dx \\
&= \operatorname{erf}(2) - \operatorname{erf}(-1) \\
&= \operatorname{erf}(2) + \operatorname{erf}(1) \\
&\simeq .82.
\end{aligned}$$

Then

$$E[X|B] = \frac{1}{.82} \frac{1}{\sqrt{2\pi}} \int_{-1}^{2} x e^{-\frac{1}{2}x^2} dx = (e^{-\frac{1}{2}} - e^{-2})/2.06 \simeq .22$$

17. For any $h(x)$, we have $E[h(X)] = \int_{-\infty}^{+\infty} h(x) f_X(x) dx$.

(a) Using Taylor's series, we write

$$\begin{aligned} h(x) &= h(\mu) + h'(\mu)(x-\mu) + h''(\mu)\frac{(x-\mu)^2}{2!} + ... \\ &\simeq h(\mu) + h'(\mu)(x-\mu) + h''(\mu)\frac{(x-\mu)^2}{2!}, \end{aligned}$$

if $|h^{(n)}(\mu)|$ is sufficiently small and $|x-\mu|$ is not too large. So, using this approximation, we have

$$\begin{aligned} E[h(X)] &= \int_{-\infty}^{+\infty} h(x) f_X(x) dx \\ &\simeq E\left[h(\mu) + h'(\mu)(X-\mu) + h''(\mu)\frac{(X-\mu)^2}{2!}\right] \\ &= E[h(\mu)] + E[h'(\mu)(X-\mu)] + E\left[h''(\mu)\frac{(X-\mu)^2}{2!}\right] \\ &= h(\mu) + h'(\mu)E[X-\mu] + \frac{h''(\mu)}{2}E[(X-\mu)^2] \\ &= h(\mu) + h'(\mu)\left(E[X]-\mu\right) + \frac{h''(\mu)}{2}\sigma^2 \\ &= h(\mu) + \frac{h''(\mu)}{2}\sigma^2. \end{aligned}$$

(b) Let $g(x) = h^2(x)$. Then using the above approximate Taylor series on function g gives:

$$\begin{aligned} g(x) &\simeq g(\mu) + g'(\mu)(x-\mu) + g''(\mu)\frac{(x-\mu)^2}{2!} \\ &= h^2(\mu) + 2h(\mu)h'(\mu)(x-\mu) + 2[h(\mu)h''(\mu) + (h'(\mu))^2]\frac{(x-\mu)^2}{2!}. \end{aligned}$$

Thus, applying the result of part a), now to the function g, we have

$$\begin{aligned} E[h^2(X)] &= E[g(X)] \\ &= g(\mu) + \frac{g''(\mu)}{2}\sigma^2 \\ &= h^2(\mu) + \frac{2[h(\mu)h''(\mu) + (h'(\mu))^2]}{2}\sigma^2 \\ &= h^2(\mu) + [h(\mu)h''(\mu) + (h'(\mu))^2]\sigma^2. \end{aligned}$$

18. (a) The joint pdf of X and Y is

$$\begin{aligned} f_{XY}(x,y) &= f_X(x) f_{Y|X}(y|x) \\ &= \begin{cases} 3x^2 \cdot 2y/x^2 = 6y, & 0 < y < x < 1, \\ 0, & \text{elsewhere.} \end{cases} \end{aligned}$$

(b) The conditional mean of Y given $X = x$ is

$$\begin{aligned} E[Y|x] &= \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy \\ &= \begin{cases} \int_0^x y\left(2y/x^2\right) dy = \frac{2}{3}x, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases} \end{aligned}$$

(c) The marginal pdf of Y is

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,y)dx \\ &= \begin{cases} \int_y^1 6y dx = 6y(1-y), & 0<y<1, \\ 0, & \text{elsewhere.} \end{cases} \end{aligned}$$

Then

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy = \int_0^1 6y^2(1-y)dy = \frac{1}{2}.$$

19. Let the number of units manufactured at the various sites be denoted n_A, n_B, and n_C, with total number of units simply n. Then from the problem statement we know that

$$n_A = 3n_B \quad \text{and} \quad n_B = 2n_C,$$

and of course $n = n_A + n_B + n_C$. Then from classical probabilities, we get the probability of a unit selected 'at random' as

$$P[A] = \frac{n_A}{n} = \frac{6}{9}, P[B] = \frac{n_B}{n} = \frac{2}{9}, \text{ and } P[C] = \frac{n_C}{n} = \frac{1}{9},$$

where we define event $A \triangleq \{\text{unit comes from plant } A\}$, and so forth for events B and C. Now we can use the concept of conditional expectation to write

$$\begin{aligned} E[X] &= E[X|A]P[A] + E[X|B]P[B] + E[X|C]P[C] \\ &= \frac{1}{5}\int_0^{\infty} xe^{-x/5}dx\frac{6}{9} + \frac{1}{6.5}\int_0^{\infty} xe^{-x/6.5}dx\frac{2}{9} + \frac{1}{10}\int_0^{\infty} xe^{-x/10}dx\frac{1}{9} \\ &= 5\frac{6}{9} + 6.5\frac{2}{9} + 10\frac{1}{9} \approx 5.89 \text{ years.} \end{aligned}$$

20. We are asked to compute the expected value $E[Y]$ of the received signal. Now

$$\begin{aligned} E[Y] &= E[E[Y|\Theta]] \\ &= \int_{-\infty}^{\infty} E[Y|\theta]f(\theta)d\theta. \end{aligned}$$

Now

$$E[Y|\Theta=\theta] = \theta \quad \text{by inspection of the given conditional Normal density.}$$

Thus

$$E[Y] = \frac{1}{2\pi}\int_0^{2\pi} \theta d\theta = \pi.$$

21. This problem computes the variances of a) Bernoulli, b) binomial, c) Poisson, d) Gaussian, and e) Rayleigh random variables.

(a) $E[X^2] = 1^2 p + 0^2(1-p) = p$, $E[X] = p$, so $\sigma^2 = E[X^2] - (E[X])^2 = p - p^2 = p(1-p)$.

(b) For the binomial random variable X, we have $E[X] = np$ from Problem 4.4. Then, letting $q = 1 - p$, we have

$$\begin{aligned}
E[X^2] &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^k q^{n-k} \\
&= \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^k q^{n-k} \\
&= \sum_{k=1}^{n} ((k-1)+1) \frac{n!}{(k-1)!(n-k)!} p^k q^{n-k} \\
&= n(n-1)p^2 + np.
\end{aligned}$$

Then $\sigma^2 = E[X^2] - (E[X])^2 = n(n-1)p^2 + np - (np)^2 = np(1-p) = npq$.

(c) From Example 4.1-6, we have $E[X] = a$ for the Poisson random variable with parameter $a(>0)$. Remember the Poisson PMF is given as $P_X(k) = \frac{a^k}{k!} e^{-a} u(k)$. To compute the second moment, we proceed

$$\begin{aligned}
E[X^2] &= \sum_{k=-\infty}^{+\infty} k^2 P_X(k) \\
&= \sum_{k=1}^{\infty} k^2 \frac{a^k}{k!} e^{-a} \qquad \text{since the } k = 0 \text{ term will be zero,} \\
&= a \left(\sum_{k=1}^{\infty} (k-1+1) \frac{a^{k-1}}{(k-1)!} \right) e^{-a} \\
&= a e^{-a} \left(\sum_{k=1}^{\infty} (k-1) \frac{a^{k-1}}{(k-1)!} + \sum_{k=1}^{\infty} 1 \frac{a^{k-1}}{(k-1)!} \right) \\
&= a e^{-a} \left(\sum_{k'=0}^{\infty} k' \frac{a^{k'}}{k'!} + \sum_{k'=0}^{\infty} 1 \frac{a^{k'}}{k'!} \right) \qquad \text{with the substitution } \ k' = k - 1, \\
&= a e^{-a} (a e^{+a} + e^{+a}) \\
&= a^2 + a.
\end{aligned}$$

Then $\sigma^2 = E[X^2] - (E[X])^2 = a^2 + a - a^2 = a(>0)$.

(d) We must compute the variance of the Gaussian random variable with parameters μ and σ^2, i.e. $X : N(\mu, \sigma^2)$. For the mean $E[X] = \mu$ see Example 4.1-2 in textbook. For the variance, consider the random variable $X - \mu$ with variance $E[(X-\mu)^2]$.

$$\begin{aligned}
E[(X-\mu)^2] &= \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z^2 e^{-\frac{z^2}{2}} dz \quad \text{with substitution } \ z \triangleq (x-\mu)/\sigma.
\end{aligned}$$

Next we integrate by parts with $u = z$ and $dv = ze^{-\frac{z^2}{2}}dz$, yielding $du = dz$ and $v = -e^{-\frac{z^2}{2}}$, so that, the above integral becomes

$$\begin{aligned}\int_{-\infty}^{+\infty} z^2 e^{-\frac{z^2}{2}} dz &= \left(-ze^{-\frac{z^2}{2}}\right)|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz \\ &= -0 + 0 + \sqrt{2\pi},\end{aligned}$$

where the last term is due to the fact that the standard normal density $N(0,1)$ integrates to 1. Thus we have $E[(X-\mu)^2] = \frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi} = \sigma^2$, and thus the parameter σ^2 in the Gaussian density is shown to be the variance of the random variable $X - \mu$, which is the same as the variance of the random variable X.

(e) We are concerned here with the Rayleigh random variable with parameters μ and σ^2. The mean is calculated as follows.

$$\begin{aligned}E[X] &= \int_0^\infty x\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}dx \\ &= \int_0^\infty \left(\frac{x}{\sigma}\right)^2 e^{-\frac{x^2}{2\sigma^2}}dx \\ &= \sigma\int_0^\infty y^2 e^{-\frac{y^2}{2}}dy \quad \text{with the transformation } y \triangleq \frac{x}{\sigma}, \\ &= \sigma\sqrt{\frac{\pi}{2}} \quad \text{from the similar } z \text{ integral done above.}\end{aligned}$$

To calculate the variance, we again rely on the indirect method of calculating $E[X^2]$ first.

We have

$$\begin{aligned}E[X^2] &= \int_0^\infty x^2\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}dx \\ &= \int_0^\infty \frac{x^3}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}dx,\end{aligned}$$

which can be calculated directly. An easier way though is to observe from Example 3.3-11 that the Raleigh random variable results from the square root of the sum of the squares of two independent Gaussians, each distributed as $N(0,\sigma^2)$. Calling this Rayleigh random variable Z, then $Z^2 = X^2 + Y^2$, and so $E[Z^2] = E[X^2 + Y^2] = E[X^2] + E[Y^2] = 2\sigma^2$. Hence the variance of the Rayleigh random variable Z with parameter σ is given via $\sigma^2 = E[Z^2] - \mu^2$, as $2\sigma^2 - \left(\sigma\sqrt{\frac{\pi}{2}}\right)^2 = \left(2 - \frac{\pi}{2}\right)\sigma^2$.

22. (a) We use conditional expectation to write

$$\begin{aligned}E[T] &= E[E[T|\text{type}]] \\ &= 0.3E[T_1] + 0.7E[T_2] \\ &= 0.3\mu_1 + 0.7\mu_2.\end{aligned}$$

(b) We again use conditional expectation to write

$$\begin{aligned}E[T^2] &= E[E[T^2|\text{type}]] \\ &= 0.3E[T_1^2] + 0.7E[T_2^2].\end{aligned}$$

Now for the exponential random variable X with mean μ_X, we remember[1] that $E[X^2] = 2\mu_X^2$. Thus we have $E[T_i^2] = 2\mu_i^2$ for $i = 1, 2$ and so then

$$\begin{aligned} E[T^2] &= 0.3E[T_1^2] + 0.7E[T_2^2] \\ &= 0.6\mu_1^2 + 1.4\mu_2^2. \end{aligned}$$

(c) We start from the general expression $\sigma^2 = E[X^2] - E^2[X]$, so that here

$$\begin{aligned} \sigma_T^2 &= E[T^2] - E^2[T] \\ &= 0.6\mu_1^2 + 1.4\mu_2^2 - (0.3\mu_1 + 0.7\mu_2)^2 \\ &= 0.51\mu_1^2 - 0.42\mu_1\mu_2 + 0.91\mu_2^2. \end{aligned}$$

Thus finally $\sigma_T = \sqrt{0.51\mu_1^2 - 0.42\mu_1\mu_2 + 0.91\mu_2^2}$.

23. Taking the indirect approach, we start with calculation of the CDF $F_Z(z)$ for $z \geq 0$ as

$$\begin{aligned} F_Z(z) &= \iint\limits_{\sqrt{x^2+y^2} \leq z} e^{-\frac{1}{2}(x^2+y^2)} dxdy \\ &= \int_0^z \int_0^{2\pi} e^{-\frac{1}{2}r^2} r dr d\theta, \\ \text{(with the transformation} &: \quad x = r\cos\theta, y = r\sin\theta, \quad \text{and} \quad dxdy = rdrd\theta), \\ &= 1 - e^{-\frac{1}{2}z^2} \quad \text{for} \quad z \geq 0. \end{aligned}$$

Since $F_Z(z) = 0$ for $z < 0$, we have the total solution $F_Z(z) = (1 - e^{-\frac{1}{2}z^2})u(z)$. Then the pdf is given as

$$\begin{aligned} f_Z(z) &= \frac{dF_Z(z)}{dz} \\ &= \frac{d[(1 - e^{-\frac{1}{2}z^2})u(z)]}{dz} \\ &= \frac{d(1 - e^{-\frac{1}{2}z^2})}{dz} u(z) + 0\delta(z) \\ &= ze^{-\frac{1}{2}z^2} u(z). \end{aligned}$$

We then calculate the mean as

$$\begin{aligned} E[Z] &= \int_{-\infty}^{+\infty} z f_Z(z) dz \\ &= \int_0^{+\infty} z^2 e^{-\frac{1}{2}z^2} dz \\ &= \sqrt{\frac{\pi}{2}}, \end{aligned}$$

[1] Alternatively, it is found using integration by parts twice.

by manipulation of the integral for the variance of a standard Normal. Then, the mean square value is

$$\begin{aligned} E[Z^2] &= \int_{-\infty}^{+\infty} z^2 f_Z(z)dz \\ &= \int_{0}^{+\infty} z^3 e^{-\frac{1}{2}z^2} dz \\ &= 2, \end{aligned}$$

by using the substitution $v \triangleq \frac{1}{2}z^2$. Finally, the variance can then be found as

$$\begin{aligned} \text{Var}[Z] &= E[Z^2] - E^2[Z] \\ &= 2 - (\sqrt{\frac{\pi}{2}})^2 \\ &= 2 - \frac{\pi}{2} \\ &\simeq 0.43. \end{aligned}$$

24. (a) Although it may not look like it, this is a standard transformation of RVs problem. First we partition the x_1, x_2, x_3space into six disjoint regions $\mathfrak{R}_1, \mathfrak{R}_2, \mathfrak{R}_3, \mathfrak{R}_4, \mathfrak{R}_5$, and $\mathfrak{R}_6$,where

$$\begin{aligned} \mathfrak{R}_1 &= \{x_1 < x_2 < x_3\}, \mathfrak{R}_2 = \{x_1 < x_3 < x_2\}, \\ \mathfrak{R}_3 &= \{x_2 < x_1 < x_3\}, \mathfrak{R}_4 = \{x_2 < x_3 < x_1\}, \\ \mathfrak{R}_5 &= \{x_3 < x_1 < x_2\}, \text{ and } \mathfrak{R}_6 = \{x_3 < x_2 < x_1\}. \end{aligned}$$

In $\mathfrak{R}_1$, write $y_1 \triangleq g_1(x_1, x_2, x_3) = x_1, y_2 \triangleq h_1(x_1, x_2, x_3) = x_2, y_3 \triangleq q_1(x_1, x_2, x_3) = x_3$.
In $\mathfrak{R}_2$, write $y_1 \triangleq g_2(x_1, x_2, x_3) = x_1, y_2 \triangleq h_2(x_1, x_2, x_3) = x_3, y_3 \triangleq q_2(x_1, x_2, x_3) = x_2$.
In $\mathfrak{R}_3$, write $y_1 \triangleq g_3(x_1, x_2, x_3) = x_2, y_2 \triangleq h_3(x_1, x_2, x_3) = x_1, y_3 \triangleq q_3(x_1, x_2, x_3) = x_3$.
In $\mathfrak{R}_4$, write $y_1 \triangleq g_4(x_1, x_2, x_3) = x_2, y_2 \triangleq h_4(x_1, x_2, x_3) = x_3, y_3 \triangleq q_4(x_1, x_2, x_3) = x_1$.
In $\mathfrak{R}_5$, write $y_1 \triangleq g_5(x_1, x_2, x_3) = x_3, y_2 \triangleq h_5(x_1, x_2, x_3) = x_1, y_3 \triangleq q_5(x_1, x_2, x_3) = x_2$.
In $\mathfrak{R}_6$, write $y_1 \triangleq g_6(x_1, x_2, x_3) = x_3, y_2 \triangleq h_6(x_1, x_2, x_3) = x_2, y_3 \triangleq q_6(x_1, x_2, x_3) = x_1$.
In each region we have a different transformation. The magnitude of the Jacobian for each transformation is unity. For example in $\mathfrak{R}_1$=$\{x_1 < x_2 < x_3\}$:

$$|J_1| = \text{mag}\begin{vmatrix} \partial g_1/\partial x_1 & \partial g_1/\partial x_2 & \partial g_1/\partial x_3 \\ \partial h_1/\partial x_1 & \partial h_1/\partial x_2 & \partial h_1/\partial x_3 \\ \partial q_1/\partial x_1 & \partial q_1/\partial x_2 & \partial q_1/\partial x_2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Computing the Jacobian for the other transformations, we get

$$|J_i| = mag\begin{vmatrix} \partial g_i/\partial x_1 & \partial g_i/\partial x_2 & \partial g_i/\partial x_3 \\ \partial h_i/\partial x_1 & \partial h_i/\partial x_2 & \partial h_i/\partial x_3 \\ \partial q_i/\partial x_1 & \partial q_i/\partial x_2 & \partial q_i/\partial x_2 \end{vmatrix} = 1, \quad i = 1, ..., 6.$$

Finally, again considering region $\mathfrak{R}_1$, we find that the only solution to $y_1 - g_1(x_1, x_2, x_3) = 0$ is $x_1^{(1)} = y_1$;$y_2 - h_1(x_1, x_2, x_3) = 0$ is $x_2^{(1)} = y_2$;$y_3 - q_1(x_1, x_2, x_3) = 0$ is $x_3^{(1)} = y_3$, where the superscript has been added to remind us that the solution applies to region

1. For example in $\Re_2$ we would have: $x_1^{(2)} = y_1; x_2^{(2)} = y_3; x_3^{(2)} = y_2$; To get the final solution we use

$$\begin{aligned} f_{Y_1Y_2Y_3}(y_1,y_2,y_3) &= \sum_{i=1}^{6} f_{X_1X_2X_3}(x_1^{(i)},x_2^{(i)},x_3^{(i)})/|J_i| \\ &= \begin{cases} 6(2\pi)^{-3/2}\exp[-(y_1^2+y_2^2+y_3^2)/2], & y_1<y_2<y_3, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

(b) To get, say, $E[Y_1]$, we need $f_{Y_1}(y_1)$ to compute $E[Y_1]=\int_{-\infty}^{\infty} y_1 f_{Y_1}(y_1)dy_1$. But $f_{Y_1}(y_1)$ is computed as

$$\begin{aligned} f_{Y_1}(y_1) &= 6\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_1^2}\frac{1}{\sqrt{2\pi}}\int_{y_3=y_1}^{\infty} e^{-\frac{1}{2}y_3^2}\left(\frac{1}{\sqrt{2\pi}}\int_{y_1}^{y_3} e^{-\frac{1}{2}y_2^2}dy_2\right)dy_3, \\ &= 6\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_1^2}\frac{1}{\sqrt{2\pi}}\int_{y_3=y_1}^{\infty} e^{-\frac{1}{2}y_3^2}\left(F_{SN}(y_3)-F_{SN}(y_1)\right)dy_3. \end{aligned}$$

To get $f_{Y_2}(y_2)$ to compute $E[Y_2]=\int_{-\infty}^{\infty} y_2 f_{Y_2}(y_2)dy_2$ we proceed as follows:

$$\begin{aligned} f_{Y_2}(y_2) &= 6\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_2^2}\frac{1}{\sqrt{2\pi}}\int_{y_2}^{\infty} e^{-\frac{1}{2}y_3^2}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y_2} e^{-\frac{1}{2}y_1^2}dy_1\right)dy_3 \\ &= 6\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_2^2}F_{SN}(y_2)\left(1-F_{SN}(y_2)\right). \end{aligned}$$

Hence $E[Y_2] = \int_{-\infty}^{\infty} y_2 f_{Y_2}(y_2)dy_2 = 0$, since the integrand is odd and the interval of integration is even.

25. The support of interest is shown below:

$$\begin{aligned} E[Y] &= 2\int_{y=0}^{1}\int_{x=0}^{y} y\,dx\,dy = 2/3, \\ E[Y^2] &= 2\int_{y=0}^{1}\int_{x=0}^{y} y^2\,dx\,dy = 1/2. \end{aligned}$$

Hence $\sigma_Y^2 = E[Y^2]-E^2[Y] = 1/2-4/9 = 1/18$.

26. We first compute the marginal density $f_Y(y)$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f_{XY}(x,y)dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2(1-\rho^2)^{1/2}}\exp\left(-\frac{x^2+y^2-2\rho xy}{2\sigma^2(1-\rho^2)}\right)dx \\ &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{y^2}{2\sigma^2}\right) \end{aligned}$$

Y is Gaussian random variable with distribution $N(0, \sigma^2)$, so $E[Y] = 0$.

$$\begin{aligned}
f_{Y|X}(y|x) &= \frac{f_{XY}(x,y)}{f_X(x)} \\
&= \frac{f_{XY}(x,y)}{\int_{-\infty}^{+\infty} f_{XY}(x,y)dy} \\
&= \frac{\frac{1}{2\pi\sigma^2\sqrt{1-\rho}^{1/2}} \exp\left(-\frac{x^2+y^2-2\rho xy}{2\sigma^2(1-\rho^2)}\right)}{\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right)} \\
&= \frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\rho^2}} \exp\left(-\frac{(y-\rho x)^2}{2\sigma^2(1-\rho^2)}\right)
\end{aligned}$$

$$\begin{aligned}
E[Y|X=x] &= \int_{-\infty}^{+\infty} y f_{Y|X}(y|x)dy \\
&= \int_{-\infty}^{+\infty} \frac{y}{\sqrt{2\pi}\sigma\sqrt{1-\rho^2}} \exp\left(-\frac{(y-\rho x)^2}{2\sigma^2(1-\rho^2)}\right) dy \\
&= \rho x
\end{aligned}$$

This result shows that although $E[Y] = 0$, the expected value of Y given X is ρX. The best predictor Y_ρ of Y is $Y_\rho = \rho X$. Of course this result only makes sense if we can observe X

27. (a) $\mu_Z = E[Z] = E[\frac{1}{2}(X+Y)] = \frac{1}{2}E[X] + \frac{1}{2}E[Y] = 0$. The variance is given by

$$\begin{aligned}
\sigma_Z^2 &= E[Z^2] - \mu_Z^2 \\
&= E\left[\left(\frac{X+Y}{2}\right)^2\right] - 0 \\
&= \frac{1}{4}E[X^2 + 2XY + Y^2] \\
&= \frac{1}{4}(\sigma^2 + 2E[XY] + \sigma^2) \\
&= \frac{1}{4}(2\sigma^2 + 0) \\
&\quad \text{(since } X, Y \text{ are independent,} E[XY] = E[X]E[Y]) \\
&= \frac{\sigma^2}{2}.
\end{aligned}$$

(b) Even if X and Y are dependent, the mean of Z would remain the same. Hence, $E[Z] = 0$. The variance of Z is given by

$$\begin{aligned}
\sigma_Z^2 &= E[\frac{1}{4}(X+Y)^2] - 0 \\
&= \frac{1}{4}E[X^2 + 2XY + Y^2] \\
&= \frac{1}{4}(\sigma^2 + 2E[XY] + \sigma^2) \\
&= \frac{1}{4}(2\sigma^2 + 2\rho\sigma^2) \\
&\quad \text{(because } E[(X-0)(Y-0)] = \rho\sigma_X\sigma_Y = \rho\sigma^2) \\
&= \frac{1}{2}\sigma^2(1+\rho).
\end{aligned}$$

(c) For $\rho = -1$, 0, and 1, the values of σ_Z^2 are 0, $\frac{\sigma^2}{2}$, and σ^2, respectively. The variance of the sample average does not reduce when the random variables are perfectly correlated. If X and Y are uncorrelated, $\rho = 0$, and since they are Gaussian, they are independent. In that case, the variance of the sample average goes down with the number of samples. For values of ρ other than 1, the variance of the sample average is less than the variance of each sample.

28. Write $f_{XY}(x,y) = f_{Y|X}(y|x)f_X(x)$. The from Example 2.6-14, we have that in the jointly Normal case,

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma^2}} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)\sigma^2}} \quad \text{and} \quad f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}.$$

Recall that

$$\lim_{a\to\infty} ae^{-\pi a^2 x^2} = \delta(x).$$

Then setting $a \triangleq 1/\sqrt{2\pi(1-\rho^2)\sigma^2}$, then $\pi a^2 = 1/(2(1-\rho^2)\sigma^2)$, and

$$f_{Y|X}(y|x) = ae^{-\pi a^2 (y-\rho x)^2},$$

thus

$$\begin{aligned} \lim_{\rho\to 1} f_{Y|X}(y|x) &= \lim_{a\to\infty} ae^{-\pi a^2 (y-\rho x)^2} \\ &= \delta(y-x). \end{aligned}$$

Hence, and also in the limit as $\rho \to 1$, we get

$$\lim_{\rho\to 1} f_{XY}(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \delta(y-x).$$

29. (a) Calculating, we obtain $P[X = -1] = P[\{\zeta = -1\}] = \frac{1}{5}$, $P[Y = 1] = P[\{\zeta = -1 \text{ or } +1\}] = P[\{\zeta = -1\}] + P[\{\zeta = +1\}] = \frac{2}{5}$, and $P[X = -1, Y = 1] = P[\{\zeta = -1\}] = \frac{1}{5}$. Then we have

$$\begin{aligned} P[X=-1]P[Y=1] &= \frac{1}{5} \cdot \frac{2}{5} \\ &\neq \frac{1}{5} \\ &= P[X=-1, Y=1]. \end{aligned}$$

Hence, by definition, X and Y are not independent RVs.

(b)

$$\begin{aligned} E[X] &= \sum xP[X=x] \\ &= \sum_{i=1}^{5} \zeta_i P[\{\zeta_i\}] \\ &= \frac{1}{5}(-1 - \frac{1}{2} + 0 + \frac{1}{2} + 1) \\ &= 0. \end{aligned}$$

$$\begin{aligned} E[Y] &= \sum yP[Y=y] \\ &= \sum_{i=1}^{5} \zeta_i^2 P[\{\zeta_i\}] \\ &= \frac{1}{5}(1+\frac{1}{4}+0+\frac{1}{4}+1) \\ &= \frac{1}{5}\frac{5}{2} = \frac{1}{2}. \end{aligned}$$

$$\begin{aligned} E[XY] &= \sum xyP[X=x,Y=y] \\ &= \sum_{i=1}^{5} \zeta_i^3 P[\{\zeta_i\}] \\ &= \frac{1}{5}(-1-\frac{1}{8}+0+\frac{1}{8}+1) \\ &= 0. \end{aligned}$$

Thus $E[XY] = 0 = 0 \cdot \frac{1}{2} = E[X]E[Y]$, and so, X and Y are uncorrelated RVs.

30. The conditional mean is always the mean of the conditional density. Since this conditional density is $N(\alpha x, \sigma^2)$, it follows that the conditional mean is αx, i.e. $E[Y|X=x] = \alpha x$, then by definition of the conditional mean as a random variable, we have

$$E[Y|X] = \alpha X.$$

31. We want to maximize

$$H(X) \triangleq -\int_{-\infty}^{+\infty} f(x)\ln f(x)dx, \qquad \text{(A)}$$

subject to the constraints:

$$\begin{aligned} \int_{-\infty}^{+\infty} xf(x)dx &= \mu, \\ \int_{-\infty}^{+\infty} x^2 f(x)dx &= \mu^2+\sigma^2, \\ \text{and } \int_{-\infty}^{+\infty} f(x)dx &= 1. \end{aligned}$$

Now, via the substitution into (A) with $y = x-\mu$, it is easily shown that H(X) is invariant with respect to the mean value μ, so that we can simplify to minimization of:

$$-H(X) = \int_{-\infty}^{+\infty} f(x)\ln f(x)dx,$$

subject to the constraints:

$$\begin{aligned} \int_{-\infty}^{+\infty} x^2 f(x)dx &= \sigma^2, \\ \text{and } \int_{-\infty}^{+\infty} f(x)dx &= 1, \end{aligned}$$

where we now assume $\mu = 0$. To solve this problem, we introduce two Lagrange multipliers, and write the augmented functional

$$g(f) = \int_{-\infty}^{+\infty} f(x) \ln f(x) dx + \lambda_1 \left(\int_{-\infty}^{+\infty} f(x) dx - 1 \right) + \lambda_2 \left(\int_{-\infty}^{+\infty} x^2 f(x) dx - \sigma^2 \right).$$

Then we form the functional derivatives

$$\begin{aligned} \frac{\partial g(f)}{\partial f} &= \ln f(x) + 1 + \lambda_1 + \lambda_2 x^2 \\ &= 0, \end{aligned}$$

which implies that

$$\begin{aligned} f(x) &= e^{-1-\lambda_1} e^{-\lambda_2 x^2} \\ &= c e^{-\lambda_2 x^2}, \end{aligned}$$

for some normalization constant c. Thus we must have a Normal random variable for maximum entropy. Bringing back the mean parameter μ, the only choice is $X : N(\mu, \sigma^2)$ which satisfies all the given constraints.

32. Write for positive integer n,

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x^n e^{-\frac{x^2}{2\sigma^2}} dx.$$

If n is odd, since the integrand is then an odd function and the integration region is even, the integral is zero. Next let

$$\alpha \triangleq \sigma^{-2} \quad \text{or} \quad \sigma = \alpha^{-\frac{1}{2}},$$

and assume that m is any positive integer, then

$$\int_{-\infty}^{+\infty} x^m e^{-\frac{1}{2}\alpha x^2} dx = \sqrt{2\pi} \alpha^{-\frac{1}{2}}.$$

Now, differentiating m times with respect to α, we get

$$\int_{-\infty}^{+\infty} \underbrace{x^2 x^2 \cdots x^2}_{m \text{ times}} (-\frac{1}{2})^m e^{-\frac{1}{2}\alpha x^2} dx = \sqrt{2\pi} \left(\underbrace{(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2}) \cdots}_{m \text{ times}} \right) \alpha^{-\frac{(2m+1)}{2}},$$

or

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x^{2m} e^{-\frac{x^2}{2\sigma^2}} dx = 1 \cdot 3 \cdot 5 \cdots (2m-1) \sigma^{2m}.$$

Then, for n even, we can substitute $n = 2m$ to obtain

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x^n e^{-\frac{x^2}{2\sigma^2}} dx = 1 \cdot 3 \cdot 5 \cdots (n-1) \sigma^n.$$

33.

$$
\begin{aligned}
& E\left[\frac{c_{11}}{c_{20}}(X-\mu_X)-(Y-\mu_Y)\right]^2 \\
= \;& \frac{c_{11}^2}{c_{20}^2}c_{20}+c_{02}-2\frac{c_{11}}{c_{20}}c_{11} \\
= \;& c_{02}-\frac{c_{11}^2}{c_{20}}
\end{aligned}
$$

But here $c_{11}^2 = c_{02}c_{20}$, so $c_{02} - \frac{c_{11}^2}{c_{20}} = c_{02} - \frac{c_{02}c_{20}}{c_{20}} = 0$. Now define a random variable

$$Z \triangleq \frac{c_{11}}{c_{20}}(X-\mu_X)-(Y-\mu_Y).$$

Then we have shown that $E[Z^2]=0$, so we can conclude that Z itself is zero by the Chebyshev inequality, i.e. $Z=0$. Then, rearranging the right-hand side of the above equation, we obtain

$$
\begin{aligned}
Y &= \frac{c_{11}}{c_{20}}X+(\mu_Y-\frac{c_{11}}{c_{20}}\mu_X) \\
&= \alpha X+\beta, \text{ with} \\
\alpha &= \frac{c_{11}}{c_{20}}, \text{ and } \beta=\mu_Y-\frac{c_{11}}{c_{20}}\mu_X.
\end{aligned}
$$

34. From the text: $\alpha_o = \rho\frac{\sigma_Y}{\sigma_X}$ and $\beta_o = \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}\mu_X$. Then

$$
\begin{aligned}
\epsilon^2_{\min} &= E[(Y-\alpha_o X-\beta_o)^2] \\
&= E[((Y-\mu_Y)-\alpha_o(X-\mu_X))^2] \\
&= \sigma_Y^2-2\rho\frac{\sigma_Y}{\sigma_X}\rho\sigma_X\sigma_Y+\left(\rho\frac{\sigma_Y}{\sigma_X}\right)^2\sigma_X^2 \\
&= \sigma_Y^2(1-\rho^2).
\end{aligned}
$$

When $\rho=1$, then Y is a linear function of X and there is no error in prediction.

35. We compute

$$
\begin{aligned}
m_r &= \int_{-\infty}^{\infty} x^r f_X(x)dx = \int_0^2 x^r(1-\frac{x}{2})dx \\
&= \int_0^2 x^r dx - \int_0^2 (x^{r+1}/2)dx \\
&= \left.\frac{x^{r+1}}{r+1}\right|_0^2 - \left.\frac{x^{r+2}}{2(r+2)}\right|_0^2.
\end{aligned}
$$

Thus $m_r = \frac{2^{r+1}}{(r+1)(r+2)}$. We find that $m_0=1, m_1=2/3, m_2=2/3, ...$

36. Computing the mean:

$$\begin{aligned} E\left[\hat{\mu}_N\right] &= E\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] \\ &= \frac{1}{N}\sum_{i=1}^{N} E[X_i] \\ &= \frac{1}{N} N\mu \\ &= \mu \end{aligned}$$

Computing the variance:

$$\begin{aligned} E\left[\hat{\mu}_N^2\right] &= E\left[\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right)^2\right] \\ &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E[X_i X_j] \\ &= \frac{1}{N^2}\left(\sum_{i=1}^{N} E\left[X_i^2\right] + \sum_{1\le i,j\le N, i\ne j} E[X_i X_j]\right) \\ &= \frac{1}{N^2}\left(N\left(\mu^2+\sigma^2\right) + N(N-1)\mu^2\right) \\ &= \frac{1}{N}\sigma^2 + \mu^2 \end{aligned}$$

then

$$\begin{aligned} \mathrm{Var}\left[\hat{\mu}_N\right] &= E\left[\hat{\mu}_N^2\right] - \left(E\left[\hat{\mu}_N\right]\right)^2 \\ &= \frac{1}{N}\sigma^2 + \mu^2 - \mu^2 \\ &= \frac{1}{N}\sigma^2 \end{aligned}$$

37. By Chebyshev inequality:

$$\begin{aligned} P\left[\left|\hat{\mu}_N - \mu\right| > 0.1\sigma\right] &\le \frac{\mathrm{Var}\left[\hat{\mu}_N\right]}{(0.1\sigma)^2} \\ &= \frac{\frac{1}{N}\sigma^2}{0.01\sigma^2} \\ &= \frac{100}{N} \end{aligned}$$

To ensure $P\left[\left|\hat{\mu}_N - \mu\right| > 0.1\sigma\right] \le 0.01$, one can set $\frac{100}{N} \le 0.01 \Leftrightarrow N \ge 10,000$.

38. a) Moment-generating function:

$$\begin{aligned} M_X(t) &= \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{tx} dx \\ &= \frac{1}{t}\left(e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}\right) \\ &= \frac{\sinh(t/2)}{t/2} \end{aligned}$$

b)

$$\begin{aligned} \frac{d}{dt}M_X(t) &= \frac{d}{dt}\left(\frac{1}{t}\left(e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}\right)\right) \\ &= \frac{1}{t}\frac{d}{dt}\left(e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}\right) + \left(e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}\right)\frac{d}{dt}\left(\frac{1}{t}\right) \\ &= \frac{1}{2t}\left(e^{\frac{1}{2}t} + e^{-\frac{1}{2}t}\right) - \frac{1}{t^2}\left(e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}\right). \end{aligned}$$

We note that $M'_X(0)$ is undefined. However, we can evaluate its limiting form as t approaches zero and use that to determine the first moment. To do this, we first substitute Taylor series expansions for the exponential terms using the well-known series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + O(x^4), \qquad \text{which converges for all finite } x.$$

It turns out that we will have to keep terms up to $(t/2)^3$ here because of the $(1/t^2)$ term in $M'_X(t)$. Thus we get the approximation

$$\begin{aligned} e^{\frac{1}{2}t} &\approx 1 + \frac{1}{2}t + \frac{\left(\frac{1}{2}t\right)^2}{2!} + \frac{\left(\frac{1}{2}t\right)^3}{3!} \\ &= 1 + \frac{1}{2}t + \frac{1}{8}t^2 + \frac{1}{48}t^3. \end{aligned}$$

with corresponding approximation for $e^{\frac{1}{2}t} + e^{-\frac{1}{2}t}$ and $e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}$ given by

$$\begin{aligned} e^{\frac{1}{2}t} + e^{-\frac{1}{2}t} &\approx 2 + \frac{1}{4}t^2 \quad \text{and} \\ e^{\frac{1}{2}t} - e^{-\frac{1}{2}t} &\approx t + \frac{1}{24}t^3. \end{aligned}$$

Finally, inserting these approximations into our expression for $M'_X(t)$, we get the following

$$\begin{aligned} M'_X(t) &= \frac{1}{2t}\left(e^{\frac{1}{2}t} + e^{-\frac{1}{2}t}\right) - \frac{1}{t^2}\left(e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}\right) \\ &\approx \frac{1}{2t}\left(2 + \frac{1}{4}t^2\right) - \frac{1}{t^2}\left(t + \frac{1}{24}t^3\right) \\ &= t^{-1} + \frac{1}{8}t \; - t^{-1} - \frac{1}{24}t \\ &= \frac{1}{12}t, \qquad \text{good near } t = 0. \end{aligned}$$

Then

$$\begin{aligned} E[X] &= \left.\frac{d}{dt}M_X(t)\right|_{t=0} \\ &= \left.\frac{1}{12}t\right|_{t=0} \\ &= 0, \end{aligned}$$

the correct value for the mean of a uniform random variable centered on 0.

39. (a)

$$\begin{aligned} M(t) &= \sum_{k=0}^{\infty} e^{tk}\frac{a^k}{k!}e^{-a} \\ &= \left(\sum_{k=0}^{\infty}\frac{(ae^t)^k}{k!}\right)e^{-a} \\ &= e^{ae^t}e^{-a} \\ &= e^{a(e^t-1)} \\ &= \exp(a(e^t-1)). \end{aligned}$$

(b)

$$\begin{aligned} \frac{dM(t)}{dt} &= \frac{de^{ae^t}}{d(ae^t)}\frac{d(ae^t)}{dt}e^{-a} \\ &= e^{ae^t}ae^te^{-a} \\ &= \left(e^{ae^t}e^{-a}\right)ae^t \\ &= M(t)ae^t. \end{aligned}$$

So, evaluating this derivative at $t = 0$, we obtain

$$\begin{aligned} \left.\frac{dM(t)}{dt}\right|_{t=0} &= M(t)ae^t|_{t=0} \\ &= M(0)ae^0 \\ &= 1a\cdot 1 \\ &= a = m_1, \text{ the mean.} \end{aligned}$$

40. There is the relevant series expansion (in *Discrete Distributions* by L.L. Johnson and S. Kotz, Wiley and Sons, 1969) whose coefficients are closely related to the negative binomial distribution. The function and its expansion are:

$$\begin{aligned} \left(Q-Pe^t\right)^{-N} &= \sum_{k=0}^{\infty}\binom{N+k-1}{N-1}\left(\frac{Pe^t}{Q}\right)^k\left(1-\frac{P}{Q}\right)^N, \quad \text{for } t \text{ small}, \\ &= \sum_{k=0}^{\infty}e^{kt}\binom{N+k-1}{N-1}\left(\frac{P}{Q}\right)^k\left(1-\frac{P}{Q}\right)^N \\ &= \sum_{k=0}^{\infty}e^{kt}P_X(k). \end{aligned}$$

Therefore

$$M_X(t) = \left(Q - Pe^t\right)^{-N}.$$

41. By definition:

$$M(t) \triangleq \left[\alpha!\beta^{\alpha+1}\right]^{-1} \int_0^\infty x^\alpha e^{-x(\frac{1}{\beta}-t)} dx.$$

From any table of *definite exponential integrals* we find that for α an integer:

$$\int_0^\infty x^\alpha e^{-x(\frac{1}{\beta}-t)} dx = \frac{\alpha!}{(1/\beta - t)^{\alpha+1}}.$$

Hence

$$\begin{aligned} M(t) &\triangleq \left[\alpha!\beta^{\alpha+1}\right]^{-1} \int_0^\infty x^\alpha e^{-x/\beta} e^{tx} dx \\ &= \left[\alpha!\beta^{\alpha+1}\right]^{-1} \frac{\alpha!}{(1/\beta - t)^{\alpha+1}} \\ &= \frac{1}{(1-\beta t)^{\alpha+1}}. \end{aligned}$$

42. We can do this simply using the MGF of the gamma distribution. From Problem 4.41 we have $M(t) = (1-\beta t)^{-(\alpha+1)}$. Hence $M'(t) = (1-\beta t)^{-(\alpha+2)}(-)(\alpha+1)(-)\beta$ and

$$\begin{aligned} M'(0) &= (\alpha+1)\beta \\ &= E[X]. \end{aligned}$$

To get the variance we compute $M''(t)|_{t=0} = E[X^2]$. We find that

$$M''(t) = -(\alpha+1)(-\beta)(-1)(\alpha+2)(-\beta)(1-\beta t)^{-(\alpha+3)}$$

so that $M''(0) = (\alpha+1)(\alpha+2)\beta^2$. Hence

$$\begin{aligned} \sigma_X^2 &= M''(0) - [M'(0)]^2 \\ &= (\alpha+1)\beta^2. \end{aligned}$$

43. The Chernoff bound on $P[X \geq a]$, where X is an exponential random variable, is given by $P[X \geq a] \leq e^{-at}M_X(t)$, where $M_X(t) \triangleq E[e^{tX}]$, the moment generating function of X is given as

$$\begin{aligned} M_X(t) &= \int_{-\infty}^\infty e^{tx} f_X(x)\, dx \\ &= \lambda \int_0^\infty e^{tx} e^{-\lambda x}\, dx \\ &= \lambda \int_0^\infty e^{-(\lambda - t)x}\, dx \text{ (will converge if } Re(t) < \lambda \text{)} \\ &= \frac{\lambda}{\lambda - t}. \end{aligned}$$

Therefore

$$P[X \geq a] \leq e^{-at} \frac{\lambda}{\lambda - t}.$$

Denoting the right-hand side as $h(t)$, we seek its minimum by setting the derivitive of h to zero. Now

$$\begin{aligned}\frac{dh}{dt} &= \frac{\lambda}{\lambda - t}(-a)e^{-at} + e^{-at}\lambda(-)\left(\frac{1}{\lambda - t}\right)^2(-) \\ &= 0,\end{aligned}$$

impies $\lambda - \frac{1}{a}$ (after a little algebra). We must note however that t must be non-negative, so that this value only holds for $a\lambda \geq 1$. Hence, for such a, we have

$$P[X \geq a] \leq e^{-a\left(\lambda - \frac{1}{a}\right)} \frac{\lambda}{\lambda - (\lambda - \frac{1}{a})} = a\lambda e^{(-a\lambda + 1)}.$$

So, the Chernoff bound is

$$\begin{aligned}P[X \geq a] &\leq a\lambda e^{-a\lambda+1} \\ &= (a\lambda e)e^{-a\lambda}, \quad \text{for} \ \ a \geq 1/\lambda.\end{aligned}$$

44. We combine both parts (a) and (b) by doing the case for general positive integer N. By the Chernoff bound

$$P[X \geq k] \leq e^{-tk}[Q - Pe^t]^{-N}.$$

Next, let $h(t) \triangleq g_1(t)[g_2(t)]^{-N}$ with $g_1(t) \triangleq e^{-tk}$ and $g_2(t) \triangleq Q - Pe^t$. Then

$$\frac{dh(t)}{dt} = g_1(-N)[g_2(t)]^{-N-1}\frac{dg_2(t)}{dt} + [g_2(t)]^{-N}\frac{dg_1(t)}{dt}.$$

Here,

$$\frac{dg_2(t)}{dt} = -Pe^t \quad \text{and} \quad \frac{dg_2(t)}{dt} = -ke^{tk}.$$

Setting the defivative $\frac{dh(t)}{dt} = 0$, then gives

$$e^{-tk}(-N)[Q - Pe^t]^{-N-1} - Pe^t + [Q - Pe^t]^{-N}(-k)e^{-tk} = 0,$$

which implies

$$e^t = \frac{kQ}{NP + kP} \quad \text{or} \quad t = \ln\left[\frac{k(Q/P)}{N+k}\right].$$

Then, recalling that $e^{\ln x} = x$, we get

$$\begin{aligned}P[X \geq k] &\leq e^{-k\ln\left[\frac{k(Q/P)}{N+k}\right]}\left[Q - P\frac{k(Q/P)}{N+k}\right]^{-N} \\ &= \left[\frac{k(Q/P)}{N+k}\right]^{-k}\left[\frac{QN}{N+k}\right]^{-N}.\end{aligned}$$

In the special case of part (a), where $N = 1$, we have

$$\begin{aligned}P[X \geq k] &\leq \left[\frac{k(Q/P)}{1+k}\right]^{-k}\left[\frac{Q}{1+k}\right]^{-1} \\ &= \frac{(k+1)}{Q}\frac{(k+1)^k}{[k(Q/P)]^k} \\ &= \frac{(k+1)^{k+1}}{Q\,[k(Q/P)]^k}.\end{aligned}$$

If also, we have $k = 1$, then

$$\begin{aligned} P[X &\geq 1] \leq \frac{(2)^2}{Q\left[2(Q/P)\right]^2} \\ &= \frac{P^2}{Q^3} = \frac{P^2}{(1+P)^3}. \end{aligned}$$

45. The characteristic function of the Cauchy random variable X, with density function $f_X(x) = \frac{\alpha}{\pi(\alpha^2+x^2)}$, is given by

$$\Phi_X(\omega) = E[e^{j\omega X}] = \int_{-\infty}^{\infty} e^{j\omega x} \frac{\alpha}{\pi(\alpha^2+x^2)}\, dx = S(-\omega),$$

where $S(\omega) \triangleq FT\left[\frac{\alpha}{\pi(\alpha^2+x^2)}\right] = \int_{-\infty}^{\infty} e^{-j\omega x} \frac{\alpha}{\pi(\alpha^2+x^2)}\, dx$.

Now consider the function $y(x) = e^{-\alpha|x|}$. The Fourier transform of this function is given by $Y(\omega) = \frac{2\alpha}{\alpha^2+\omega^2}$. Therefore, $FT\left[\frac{1}{2\pi} y(x)\right] = \frac{1}{2\pi} Y(\omega) = \frac{\alpha}{\pi(\alpha^2+\omega^2)}$.

From the duality of the Fourier transform, we know that if $FT\{y(x)\} = Y(\omega)$, then $FT\left[Y(x)\right] = 2\pi y(-\omega)$. Therefore, $S(\omega) = FT\left[\frac{\alpha}{\pi(\alpha^2+x^2)}\right] = 2\pi \frac{1}{2\pi} y(-\omega) = y(-\omega) = e^{-\alpha|\omega|}$.

Therefore,

$$\Phi_X(\omega) = S(-\omega) = e^{-\alpha|\omega|} \quad \text{for all} \quad \omega.$$

46. Since $f_X(x) = \frac{1}{\pi(1+(x-a)^2)}$ is a pdf for any value of a including $a = 0$, *we* immediately deduce that $\int_{-\infty}^{\infty} \frac{1}{\pi(1+(x-a)^2)} dx = 1$. Then

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} \frac{x}{\pi(1+(x-a)^2)} dx \\ &= \int_{-\infty}^{\infty} \frac{u+a}{\pi(1+u^2)} du \\ &= \int_{-\infty}^{\infty} \frac{u}{\pi(1+u^2)} du + \int_{-\infty}^{\infty} \frac{a}{\pi(1+u^2)} du \\ &= 0 + a = a. \end{aligned}$$

We took advantage that the first integral after the third equal sign is 0 because the integrand is odd and the interval about zero is even. To calculate the variance we would have to consider an integral of the form $E[X^2] = \int_{-\infty}^{\infty} \frac{x^2}{\pi(1+(x-a)^2)} dx$, which is readily seen not to converge in any sense. See what happens to the integrand when x approaches infinity.

47. The solution proceeds as

$$\begin{aligned} \Phi_X(\omega) &\triangleq E[e^{+j\omega X}] \\ &= \int_0^{\infty} \frac{1}{\mu} e^{(j\omega - 1/\mu)x} dx \\ &= \frac{1}{\mu} \left. \frac{e^{(j\omega-1/\mu)x}}{j\omega - \frac{1}{\mu}} \right|_0^{\infty} \\ &= \frac{1}{\mu} \frac{-1}{j\omega - \frac{1}{\mu}} \\ &= \frac{1}{1 - j\omega\mu}, \qquad -\infty < \omega < +\infty. \end{aligned}$$

48. Calculating the CF when mean parameter $\alpha = \mu = 0$, when X is Cauchy distributed, we get $\Phi_X(\omega) \triangleq E[e^{+j\omega X}] = e^{-|\omega|}$ by a known Fourier transform relation. It is easily checked by finding the pdf that corresponds to this CF as follows:

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \Phi_X(\omega) e^{-j\omega x} d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-|\omega|} e^{-j\omega x} d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{0} e^{+\omega} e^{-j\omega x} d\omega + \frac{1}{2\pi}\int_{0}^{+\infty} e^{-\omega} e^{-j\omega x} d\omega \\ &= \frac{1}{\pi\left(1+x^2\right)}. \end{aligned}$$

So for the given case, with mean μ, we have $X_i = X + \mu$, and so its CF is given as

$$\begin{aligned} \Phi_{X_i}(\omega) &\triangleq E[e^{+j\omega X_i}] \\ &= E[e^{+j\omega(X_i+\mu)}] \\ &= e^{j\omega\mu} e^{-|\omega|}. \end{aligned}$$

Now turning to the sum Y, we have $Y = \frac{1}{N}\sum_{i=1}^{N} X_i$. Then, since the X_i are i.i.d., the CF of Y is given as

$$\begin{aligned} \Phi_Y(\omega) &= E[e^{+j\omega Y}] \\ &= \prod_{i=1}^{N} \int_{-\infty}^{+\infty} e^{+j\frac{\omega}{N}x_i} \frac{1}{\pi[1+(x_i-\mu)^2}dx_i \\ &= \prod_{i=1}^{N} e^{+j\frac{\omega}{N}\mu} e^{-\frac{1}{N}|\omega|}, \quad \text{with the substitution:} \quad z_i = x_i - \mu, \\ &= e^{+j\omega\mu - |\omega|}, \end{aligned}$$

which is equal to the $\Phi_{X_i}(\omega)$, which are all equal since the X_i are i.i.d. So, since $\Phi_Y(\omega) = \Phi_{X_i}(\omega)$, it must be that $f_Y(y) = f_{X_i}(y)$, as was to be shown. We conclude that the average of i.i.d. Cauchy RVs is also Cauchy, and with the same mean μ.

49. If X is uniform in (a, b), then $E[X] = \frac{a+b}{2}$ (see problem 4.5).

$$E[Z] = E[X+Y] = E[X] + E[Y] = \frac{a-a}{2} + \frac{na+(n-2)a}{2} = (n-1)a.$$

Since the convolution of two *rect* functions is a triangle function of twice the width with the center at the mean, the density function of Z, obtained by convolving the uniform density functions[2] of X and Y, is obtained as seen in Fig. 1.

Equivalently, $\Phi_X(\omega) = \frac{sin(\omega a)}{\omega a}$; $\Phi_Y(\omega) = \frac{sin(\omega a)}{\omega a} e^{-j\omega a(n-1)}$.

$$f_Z(z) = FT^{-1}\left[\left(\frac{sin(\omega a)}{\omega a}\right)^2 e^{-j\omega a(n-1)}\right],$$

[2] Note that the uniform density function is merely a shifted and scaled rect. function.

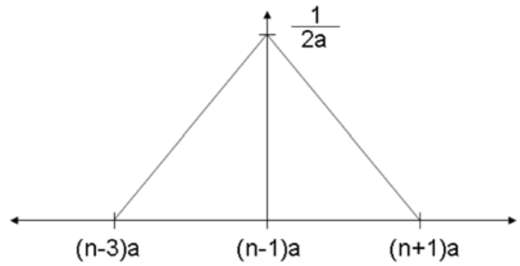


Figure 1:

which is the triangle function, merely shifted by $(n-1)a$, and results in a shift in the density function.

50.

51.

52. With $Z = aX + bY$ and $W = cX + dY$ we can write:

$$\begin{aligned}\Phi_{ZW}(\omega_1, \omega_2) &= E\left[e^{j(\omega_1 Z + \omega_2 W)}\right] \\ &= E\left[e^{j(\omega_1(aX+bY)+\omega_2(cX+dY))}\right] \\ &= E\left[e^{j((a\omega_1+c\omega_2)Z+(b\omega_1+d\omega_2)W)}\right] \\ &= \Phi_{XY}(a\omega_1 + c\omega_2, b\omega_1 + d\omega_2)\end{aligned}$$

53. We note that X and Y are independent since

$$\begin{aligned}f_{XY}(x, y) &= 16e^{-4(x+y)}u(x)u(y) \\ &= 4e^{-4x}u(x)4e^{-4y}u(y).\end{aligned}$$

Hence the joint MGF is

$$M_{XY}(t_1, t_2) = M_X(t_1)M_Y(t_2)$$

where

$$\begin{aligned}M_X(t_1) &= 4\int_0^\infty e^{-4x}e^{t_1 x}dx \\ &= 4\int_0^\infty e^{-(4-t_1)xt_1 x}dx \\ &= \frac{1}{1 - t_1/4}\end{aligned}$$

and

$$\begin{aligned}M_Y(t_2) &= 4\int_0^\infty e^{-4y}e^{t_2 y}dy \\ &= 4\int_0^\infty e^{-(4-t_2)y}dy \\ &= \frac{1}{1 - t_2/4}.\end{aligned}$$

Hence

$$\begin{aligned} M_{XY}(t_1, t_2) &= (1 - t_1/4)^{-1}(1 - t_2/4)^{-1} \\ &= \frac{16}{(4 - t_1)(4 - t_2)}. \end{aligned}$$

To get the CF, just replace t_i by $j\omega_i$ on the right-hand side. We get

$$\begin{aligned} \Phi_{XY}(\omega_1, \omega_2) &= (1 - j\omega_1/4)^{-1}(1 - j\omega_2/4)^{-1} \\ &= \frac{16}{(4 - j\omega_1)(4 - j\omega_2)}. \end{aligned}$$

54. Two random variables X: Poisson(a) and Y: Poisson (b) are given, where $a = 2$, and $b = 3$, $Z = X + Y$, and the characteristic function of Z is given by

$$\begin{aligned} \Phi_Z(\omega) &= E[e^{j\omega Z}] \\ &= E[e^{j\omega(X+Y)}] \\ &= E[e^{j\omega X}e^{j\omega Y}] \\ &= E[e^{j\omega X}]E[e^{j\omega Y}] \text{ (because } X \text{ and } Y \text{ are independent)} \\ &= \Phi_X(\omega)\Phi_Y(\omega). \end{aligned}$$

The characteristic function of X is given by

$$\begin{aligned} \Phi_X(\omega) &= E[e^{j\omega X}] \\ &= \sum_{k=0}^{\infty} e^{j\omega k} P_X(k) \\ &= \sum_{k=0}^{\infty} e^{j\omega k} \frac{e^{-a}a^k}{k!} \\ &= e^{-a} \sum_{k=0}^{\infty} \frac{(ae^{j\omega})^k}{k!} \\ &= e^{-a} e^{ae^{j\omega}} \\ &= \exp(-a + ae^{j\omega}) \\ &= \exp(-a(1 - e^{j\omega})). \end{aligned}$$

Similarly, we obtain $\Phi_Y(\omega) = \exp(-b(1 - e^{j\omega}))$. Therefore,

$$\begin{aligned} \Phi_Z(\omega) &= \exp(-a(1 - e^{j\omega}))\exp(-b(1 - e^{j\omega})) \\ &= \exp(-(a + b)(1 - e^{j\omega})). \end{aligned}$$

This implies that Z is Poisson$(a + b)$ and has a density function $f_Z(z) = \frac{e^{-5}5^n}{n!}$, for $n = 0, 1, 2, \ldots$, for $a = 2$ and $b = 3$.

55. We use the central limit theorem (CLT)

Let $X_i, i = 1, \ldots, 2000$ denote the state of the ith toaster

$$X_i = \begin{cases} 1 & \text{if toaster dented with probability } p \\ 0 & \text{if toaster OK with probability } q = 1 - p \end{cases}$$

Let $W \triangleq \sum_{i=1}^{n} X_i$, then

$$E[W] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = np$$

$$\begin{aligned}
\text{Var}[W] &= E[W^2] - (E[W])^2 \\
&= E\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left(\sum_{i=1}^{n} E[X_i]\right)^2 \\
&= E\left[\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} X_i\, X_j\right] - \left(\sum_{i=1}^{n} E[X_i]\right)^2 \\
&= \sum_{i=1}^{n} E\left[X_i^2\right] + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} E\left[X_i X_j\right] - \left(\sum_{i=1}^{n} E[X_i]\right)^2 \\
&= np + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} E[X_i]E[X_j] - n^2p^2 X_i \;\text{ and }\; X_j \text{ are independent} \\
&= np + n(n-1)p^2 - n^2p^2 \\
&= npq
\end{aligned}$$

Therefore $E[W] = np = 2000 \times 0.05 = 100$, $\text{Var}[W] = npq = 2000 \times 0.05 \times 0.95 = 95$.

$$\begin{aligned}
P[110 \le W \le 2000] &= \frac{1}{\sqrt{2\pi \times 95}} \int_{110}^{2000} e^{-\frac{1}{2}\left[\frac{x-100}{\sqrt{95}}\right]^2} dx \\
&\approx \frac{1}{\sqrt{2\pi}} \int_{1.03}^{+\infty} e^{-\frac{1}{2}x^2} dx \\
&= \frac{1}{2} - \text{erf}(1.03) \\
&= 0.15
\end{aligned}$$

56. Now

$$\begin{aligned}
P[L \le x] &= F_L(x) \\
&= 1 - e^{-0.002x} \\
&= 1 - e^{-x/\mu.}
\end{aligned}$$

Hence, $\mu = 1/0.002 = 500$.

(a)

$$E[L] = 500, \text{ and } E\left[\sum L_i\right] = 400 \cdot 500 = 200K\text{bytes}.$$

(b)

$$\sigma_X = 1/0.002 = 500 \quad \text{and } n = 400,$$

so

$$\begin{aligned}
P\left[\sum L_i > 420\right] &= 1 - \Phi\left(\frac{520-500}{(500/20)}\right) \\
&= 1 - \Phi(0.80) \\
&= 0.21.
\end{aligned}$$

57. We have 100 independent and i.i.d. random variables, say X_i, with means μ and variances σ^2. We form their sample mean

$$\widehat{\mu}_{100} = \frac{1}{100}\sum_{i=1}^{100} X_i.$$

Since the X_i are independent, we have the mean and variance of the sample mean given by

$$\mu_{\widehat{\mu}_{100}} = \mu \quad \text{and} \quad \sigma^2_{\widehat{\mu}_{100}} = \frac{1}{100}\sigma^2.$$

We can now write Chebyshev's inequality for this sample mean random variable $\widehat{\mu}_{100}$ as

$$P[|\widehat{\mu}_{100\cdot} - \mu| > \delta] \le \frac{\sigma^2_{\widehat{\mu}_{100}}}{\delta^2}.$$

Setting $\delta = \sigma/5$, we obtain

$$\begin{aligned} P[|\widehat{\mu}_{100\cdot} - \mu| > \sigma/5] &\le \frac{\sigma^2_{\widehat{\mu}_{100}}}{(\sigma/5)^2} \\ &= \frac{1}{100}\sigma^2\frac{25}{\sigma^2} \\ &= \frac{1}{4}. \end{aligned}$$

58. We use the central limit theorem (CLT)

Let $X_i, i = 1, \ldots, 2000$ denote the state of the ith panel

$$X_i = \begin{cases} 1 & \text{if panel is bad, with probability } p \\ 0 & \text{if panel is good, with probability } q = 1 - p \end{cases}$$

Let $W \triangleq \sum_{i=1}^{n} X_i$, then

$$E[W] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = np$$

$$\begin{aligned} \text{Var}[W] &= E[W^2] - (E[W])^2 \\ &= E\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left(\sum_{i=1}^{n} E[X_i]\right)^2 \\ &= E\left[\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} X_i\,X_j\right] - \left(\sum_{i=1}^{n} E[X_i]\right)^2 \\ &= \sum_{i=1}^{n} E\left[X_i^2\right] + \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} E\left[X_iX_j\right] - \left(\sum_{i=1}^{n} E[X_i]\right)^2 \\ &= np + \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} E[X_i]E[X_j] - n^2p^2 X_i \text{ and } X_j \text{ are independent} \\ &= np + n(n-1)p^2 - n^2p^2 \\ &= npq \end{aligned}$$

Therefore $E[W] = np = 2000 \times 0.03 = 60$, $\text{Var}[W] = npq = 2000 \times 0.03 \times 0.97 = 58.2$.

$$\begin{aligned} P[71 \le W \le 2000] &= \frac{1}{\sqrt{2\pi \times 58.2}} \int_{71}^{2000} e^{-\frac{1}{2}\left[\frac{x-60}{\sqrt{58.2}}\right]^2} dx \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{1.44}^{+\infty} e^{-\frac{1}{2}x^2} dx \\ &= \frac{1}{2} - \text{erf}(1.44) \\ &= 0.074 \end{aligned}$$

59. (a)

$$\begin{aligned} E[Z_n] &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} E[W_i] \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{E[X_i] - p}{\sqrt{pq}} \\ &= 0 \qquad \text{since } E[X_i] = p \end{aligned}$$

(b) The Poisson approximation is obtained by setting $a \triangleq np$ and using $p(k_1) = e^{-a}\frac{a^{k_1}}{k_1!}$ The CLT uses the following. Let X be the number of successes. Then with $W \triangleq \frac{X-np}{\sqrt{npq}}$, $E[W] = 0, \text{Var}[W] = 1$, and

$$b[k_1; n, p] \approx P[c_L \le W \le c_U]$$

where

$$\begin{aligned} c_U &\triangleq (k_1 + \frac{1}{2} - np)/\sqrt{npq} \\ c_L &\triangleq (k_1 - \frac{1}{2} - np)/\sqrt{npq} \end{aligned}$$

and $W : N(0,1)$. Note that X is the sum of a large number (n) of Bernoulli RVs, which is the basic for using the CLT.

The 3 mini-programs are summarized into one MATLAB function called "CalcBinProb(p, k1, n)". The test cases for $n = 2000, p = 0.05, k_1 = \{110, 120, 150, 170\}$ are included in MATLAB script "problem_4_59.m". The results in table 1 were obtained from the test:

Table 1: $P[k$ successes in n tries$]$

	$k_1 = 110$	$k_1 = 120$	$k_1 = 150$	$k_1 = 170$
EXACT	0.0235	4.7×10^{-4}	3.5×10^{-7}	1.45×10^{-11}
POISSON	0.0234	5.8×10^{-4}	6.5×10^{-7}	5.1×10^{-11}
CLT	$\underbrace{0.0242}_{1\sigma\text{ error}}$	$\underbrace{3.6 \times 10^{-4}}_{3\sigma\text{ error}}$	$\underbrace{8.0 \times 10^{-8}}_{5\sigma\text{ error}}$	$\underbrace{2.6 \times 10^{-13}}_{10\sigma\text{ error}}$

60. These are repeated Bernoulli trials resulting in the binomial distribution with $n = 1000$ and $p = 0.001$. Let X_i be the individual random variables, taking on value 1 for an erroneous line and 0 for an error-free line. Then we can write the sum or total of the errors as

$$Z = \sum_{i=1}^{n} X_i.$$

Then Z is Binomial with $\mu_Z = np = 1$ and $\sigma_Z^2 = npq = 0.999$. We can use the Poisson approximation to the Binomial with $a = np = 1$ here. Then

$$\begin{aligned} P[2 &\leq Z \leq 1000] = 1 - P_Z(0) - P_Z(1) \\ &\approx 1 - e^{-a} - ae^{-a} \\ &= 1 - 2e^{-1} \\ &\doteq 0.264. \end{aligned}$$

The CLT approximation gives a Normal distribution with mean $\mu = 1$ and $\sigma = \sqrt{0.999} = 0.9995$. However, it is not as accurate here since the mean μ_Z is only approximately one standard deviation away from 0, the minimum value of a Bernoulli random variable. Calculating the CLT approximate answer, we find

$$\begin{aligned} P[2 \leq Z \leq 1000] &\approx \frac{1}{\sqrt{2\pi \times 0.999}} \int_2^{1000} e^{-\frac{1}{2}\left[\frac{z-1}{\sqrt{0.999}}\right]^2} dz \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{1.0005}^{+\infty} e^{-\frac{1}{2}x^2} dx \\ &= 0.5 - erf\,(1.0005) \\ &\doteq 0.5 - 0.341 \\ &= 0.159, \quad \text{not very accurate here.} \end{aligned}$$

61. The pdf of each random variable is given as shown in Fig. **??**.

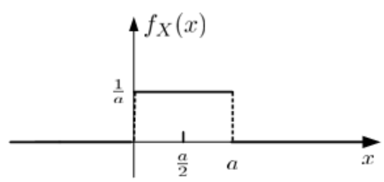


Figure 2:

(a). $E[Z] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = nE[X_1]$ since the X_i are i.i.d.

$$E[X_1] = \frac{1}{a} \int_0^a x dx = \frac{a}{2}$$

Therefore $E[Z] = \frac{na}{2}$.

(b) Define $X_i' \triangleq X_i - \frac{a}{2}$, then the X_i' are i.i.d. and $\text{Var}[X_i'] = \text{Var}[X_i]$ since shift does not affect variance.

$$\text{Var}[X_i'] = \frac{1}{a}\int_{-\frac{a}{2}}^{\frac{a}{2}} x^2 dx = \frac{x^3}{3a}\Big|_{-\frac{a}{2}}^{\frac{a}{2}} = \frac{a^2}{12}$$

Since X_i' are i.i.d.,

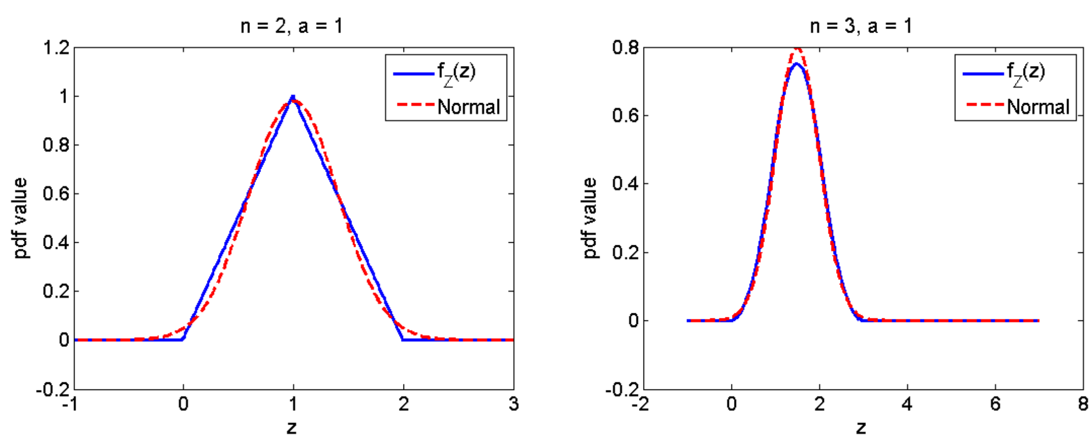
$$\begin{aligned}\text{Var}[\sum_{i=1}^{n} X_i'] &= \sum_{i=1}^{n}\text{Var}(X_i)\\ &= \frac{na^2}{12}.\end{aligned}$$

(c)(d). Characteristic function of X_i is

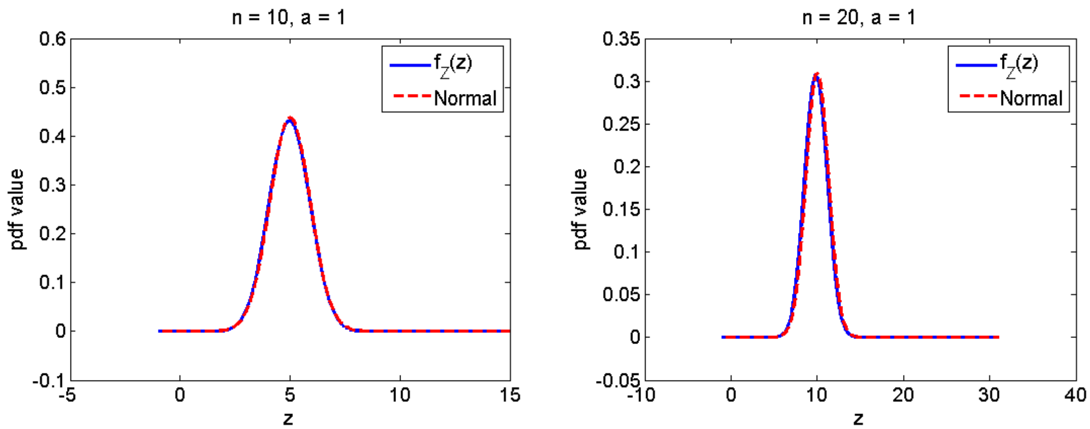
$$\begin{aligned}\Phi_{X_i}(\omega) &= \int_{-\infty}^{\infty}\frac{1}{a}e^{j\omega x}\text{rect}\left(\frac{x-a/2}{a}\right)dx\\ &= e^{\frac{j\omega a}{2}}\frac{\sin(\omega a/2)}{\omega a/2}\\ \Phi_{Z_n}(\omega) &= [\Phi_{X_i}(\omega)]^n\\ &= e^{\frac{jn\omega a}{2}}\left[\frac{\sin(\omega a/2)}{\omega a/2}\right]^n\end{aligned}$$

To get $f_{Z_n}(z)$, we can apply FFT to $\Phi_{Z_n}(\omega)$.

The MATLAB function to compute f_{Z_n} and the interval probability $P[\mu_n - k\sigma_n \le Z_n \le \mu_n + k\sigma_n]$ is called `CltVSNorm(n, a, L, Ks)`. You can run MATLAB script `problem_4_59.m` to call this function for different n. The four figures below (together called Figure 2) compare $f_{Z_n}(z)$ with Gaussian pdf's $N(\frac{na}{2}, \frac{na^2}{12})$ for $n = 2, 3, 10, 20$.



Plots of $f_Z(z)$ and its Normal approximation



(e). The probability $P(\mu - k\sigma \leq Z \leq \mu + k\sigma)$ with different n and k are shown in table 2.

Table 2: Interval probabilities $P(\mu - k\sigma \leq Z \leq \mu + k\sigma)$ in problem 4.39

	$k = 0.1$	$k = 0.5$	$k = 1$	$k = 2$	$k = 3$
$n = 2$	0.0766	0.3650	0.6475	0.9677	1.0000
$n = 3$	0.0702	0.3646	0.6667	0.9583	1.0000
$n = 10$	0.0806	0.3754	0.6783	0.9559	0.9982
$n = 20$	0.0809	0.3803	0.6770	0.9536	0.9976
Normal approximation	0.0797	0.3829	0.6827	0.9545	0.9973

62.

63.

64.

65. The MATLAB program for this problem is included in file `problem_4_65.m`. For the $\chi^2 : Z = \sum_{i=1}^{n} X_i^2$, where $X_i : N(0,1)$ are i.i.d., then

$$\begin{aligned}
\Phi_Z(\omega) &= E[e^{j\omega Z}] \\
&= E[e^{j\omega \sum_{i=1}^{n} X_i^2}] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{i=1}^{n} \left(e^{j\omega x_i^2} e^{-\frac{1}{2}x_i^2} \right) dx_1 \ldots dx_n \\
&= \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2(1-2j\omega)} dx \right]^n \\
&= \left[\frac{1}{\sqrt{1-2j\omega}} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\beta^2} d\beta}_{=1} \right]^n \qquad \beta = x\sqrt{1-2j\omega} \\
&= (1-2j\omega)^{-\frac{n}{2}}
\end{aligned}$$

(a). The graphs of the Chi-square distribution for $n = 30, 40, 50$ are shown in figure 3.

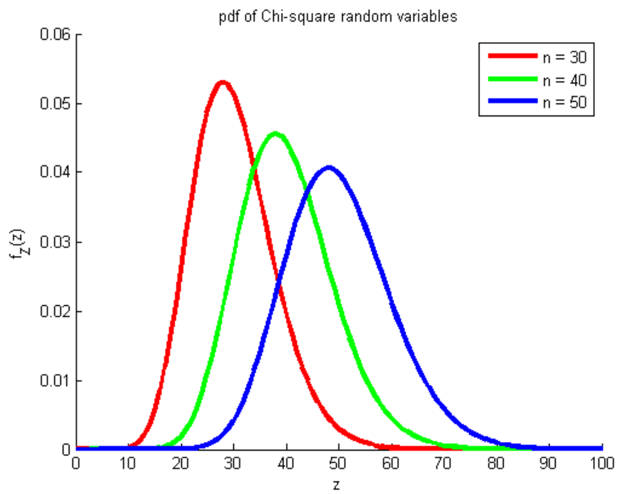


Figure 3: Graphs of Chi-square distribution

(b). Table 3 shows the probability of $P[\mu-\sigma \leq Z_n \leq \mu+\sigma]$ with $n = 30, 40, 50$ and Gaussian approximation.

Table 3: Interval probabilities $P(\mu - \sigma \leq Z_n \leq \mu + \sigma)$ in problem 4.43

$n = 30$	$n = 40$	$n = 50$	$N(n, 2n)$
0.6726	0.6620	0.6610	0.6827

66.

67. For the $\chi^2 : Z = \sum_{i=1}^{n} X_i^2$, where $X_i : N(0,1)$ are i.i.d. Then

$$
\begin{aligned}
\Phi_Z(\omega) &= E[e^{j\omega Z}] \\
&= E[e^{j\omega \sum_{i=1}^{n} X_i^2}] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{i=1}^{n} \left(e^{j\omega x_i^2} e^{-\frac{1}{2}x_i^2} \right) dx_1 \ldots dx_n \\
&= \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2(1-2j\omega)} dx \right]^n \\
&= \left[\frac{1}{\sqrt{1-2j\omega}} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\beta^2} d\beta}_{=1} \right]^n \qquad (\beta = x\sqrt{1-2j\omega}) \\
&= (1-2j\omega)^{-\frac{n}{2}}
\end{aligned}
$$

68. Let $E[X_i] = \mu$, $\text{Var}[X_i] = \sigma^2$. We wish to estimate μ with the *sample mean*

$$\widehat{\mu} \triangleq \frac{1}{N}\sum_{i=1}^{N} X_i.$$

The mean of $\widehat{\mu}$:

$$E[\widehat{\mu}] = E\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N}\sum_{i=1}^{N} E[X_i] = \frac{1}{N}\sum_{i=1}^{N} \mu = \mu$$

The variance of $\widehat{\mu}$:

$$\begin{aligned}
\text{Var}\left[\widehat{\mu}\right] &= E\left[\left(\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right)^2\right] \\
&= \frac{1}{N^2} E\left[\sum_{i=1}^{N}(X_i - \mu)\right]^2 \\
&= \frac{1}{N^2} E\left[\sum_{i=1}^{N}\sum_{j=1}^{N}(X_i - \mu)(X_j - \mu)\right] \\
&= \frac{1}{N^2} \sum_{i=1}^{N}\sum_{j=1}^{N} E\left[(X_i - \mu)(X_j - \mu)\right] \\
&= \frac{1}{N^2} \left[\sum_{i=1}^{N} E\left[(X_i - \mu)^2\right] + \sum_{i=1}^{N}\sum_{j\neq i} E\left[(X_i - \mu)(X_j - \mu)\right]\right]
\end{aligned}$$

Since X_i's are independent, so they are uncorrelated, which means $E\left[(X_i - \mu)(X_j - \mu)\right] = 0$ for all $i \neq j$. Besides, we have $E\left[(X_i - \mu)^2\right] = \text{Var}[X_i] = \sigma^2$. Thus,

$$\text{Var}\left[\widehat{\mu}\right] = \frac{1}{N^2}\sum_{i=1}^{N} \sigma^2 = \frac{\sigma^2}{N}.$$

The second moment is then

$$\begin{aligned}
E[\widehat{\mu}^2] &= \text{Var}[\widehat{\mu}] + (E[\widehat{\mu}])^2 \\
&= \frac{\sigma^2}{N} + \mu^2.
\end{aligned}$$

69.

70.

71.

72. (a) Since $Y = X + N$, the conditional PMF of Y given $X = x$,i.e. given that X is known, is

$$P_{Y|X}(y|x) = P_N(y - x).$$

This is true because given the condition $X = x$, then the equation for Y is $Y = x + N$, a simple translation by the scalar x, with the one solution $N = Y - x$. A more complicated way to see this is to start with the random variable pair (X, N) and consider the transformation to the pair (X, Y). We have $x = g(x) = x$ and $y = h(x, n) = x + n$ with inverse transformation $x = \phi(x) = x$ and $n = \psi(x, y) = y - x$. So

$$\begin{aligned} P_{X,Y}(x, y) &= P_{X,N}(x, y - x) \\ &= P_X(x)P_N(y - x), \end{aligned}$$

since X and N are independent, thus

$$\begin{aligned} P_{Y|X}(y|x) &= \frac{P_{X,Y}(x, y)}{P_X(x)} \\ &= \frac{P_X(x)P_N(y - x)}{P_X(x)} \\ &= P_N(y - x). \end{aligned}$$

Now N is Poisson with $\mu = 5$, then

$$P_N(n) = \frac{5^n}{n!}e^{-5}u(n),$$

and thus finally

$$\begin{aligned} P_{Y|X}(y|x) &= P_N(y - x) \\ &= \frac{5^{y-x}}{(y - x)!}e^{-5}u(y - x), \end{aligned}$$

with support on $y \geq x$ as it should be.

(b) Since the conditional PMF $P_{Y|X}(y|x)$ found in part a is just the Poisson PMF $P_N(n)$ shifted right by x, then the conditional mean $E[Y|X = x]$ must be $\mu + x = x + 5$. Alternatively, we can directly use the linearity of conditional expectation, as follows

$$\begin{aligned} E[Y|X = x] &= E[X + N|X = x] \\ &= E[X|X = x] + E[N|X = x] \\ &= x + E[N] \\ &= x + 5, \end{aligned}$$

where the second to last line follows since X and N are independent.

73.

74. (a) First, we consider the special case $E[X] = \mu_X = 0$.

$$\begin{aligned}
E[X^2] &= \sigma_X^2 \\
E[XY] &= \mathrm{Cov}[X,Y] + \mu_X\mu_Y = \rho\sigma_X\sigma_Y \\
\mathrm{Cov}[\varepsilon, X] &= E[\varepsilon X] - E[\varepsilon]E[X] \\
&= E[(\alpha X + \beta - Y)X] - 0 \\
&= \alpha E[X^2] + \beta E[X] - E[XY] \\
&= \alpha\sigma_X^2 - \rho\sigma_X\sigma_Y.
\end{aligned}$$

When $\mu_X \neq 0$, define $X' = X - \mu_x, \varepsilon' = \alpha X' + \beta - Y$, we can easily get $\mu_{X'} = 0, \sigma_{X'} = \sigma_X, \rho_{X'Y} = \rho_{XY} = \rho, X' - \mu_{X'} = X - \mu_X, \varepsilon' - E[\varepsilon'] = \varepsilon - E[\varepsilon]$.

$$\begin{aligned}
\mathrm{Cov}[\varepsilon, X] &= E[(\varepsilon - E[\varepsilon])(X - \mu_X)] \\
&= E[(\varepsilon' - E[\varepsilon'])(X' - E[X'])] \\
&= \mathrm{Cov}[\varepsilon', X'] \\
&= \alpha\sigma_{X'}^2 - \rho_{X'Y}\sigma_{X'}\sigma_Y \\
&= \alpha\sigma_X^2 - \rho\sigma_X\sigma_Y.
\end{aligned}$$

When $\mu_X \neq 0$, we can also start from $\mathrm{Cov}[\varepsilon, X] = E[\varepsilon X] - E[\varepsilon]E[X]$ to get the answer.

$$\begin{aligned}
\mathrm{Cov}[\varepsilon, X] &= E[\varepsilon X] - E[\varepsilon]E[X] \\
&= E[(\alpha X + \beta - Y)X] - \mu_X(\alpha\mu_X + \beta - \mu_Y) \\
&= \alpha E[X^2] + \beta\mu_X - E[XY] - \mu_X E[\varepsilon]
\end{aligned}$$

Since $E[X^2] = \sigma_X^2 + \mu_X^2, E[XY] = \rho\sigma_X\sigma_Y + \mu_X\mu_Y$, and $E[\varepsilon] = E[\alpha X + \beta - Y] = \alpha\mu_X + \beta - \mu_Y$,

(b) Set α and β to their optimal values. Then evaluate $\mathrm{Cov}[\varepsilon, X]$ again. To minimize the mean-square error $E[\varepsilon^2]$, we solve α and β that satisfy $\frac{\partial E[\varepsilon^2]}{\partial\alpha} = 0$ and $\frac{\partial E[\varepsilon^2]}{\partial\beta} = 0$.

$$\begin{aligned}
\frac{\partial E[\varepsilon^2]}{\partial\beta} &= E\left[\frac{\partial\varepsilon^2}{\partial\beta}\right] = E\left[2\varepsilon\frac{\partial\varepsilon}{\partial\beta}\right] \\
&= 2E[\varepsilon] = 2(\alpha\mu_X + \beta - \mu_Y) = 0 \\
\Rightarrow \beta &= \mu_Y - \alpha\mu_X
\end{aligned}$$

Hence, $\varepsilon = \alpha X + \mu_Y - \alpha\mu_X - Y = \alpha(X - \mu_X) - (Y - \mu_Y)$.

$$\begin{aligned}
\frac{\partial E[\varepsilon^2]}{\partial\alpha} &= E\left[\frac{\partial\varepsilon^2}{\partial\alpha}\right] = E\left[2\varepsilon\frac{\partial\varepsilon}{\partial\alpha}\right] \\
&= 2E[\varepsilon(X - \mu_X)] = 0 \\
E[\varepsilon(X - \mu_X)] &= E\left\{[\alpha(X - \mu_X) - (Y - \mu_Y)](X - \mu_X)\right\} \\
&= \alpha E[(X - \mu_X)^2] - E[(X - \mu_X)(Y - \mu_Y)] \\
&= \alpha\sigma_X^2 - \rho\sigma_X\sigma_Y = 0 \\
\Rightarrow \alpha &= \frac{\rho\sigma_Y}{\sigma_X} \\
\beta &= \mu_Y - \frac{\rho\sigma_Y\mu_X}{\sigma_X}
\end{aligned}$$

With the optimal α and β are employed, then we obtain $\text{Cov}[\varepsilon, X] = \frac{\rho\sigma_Y}{\sigma_X}\sigma_X^2 - \rho\sigma_X\sigma_Y = 0$. That is to say, the estimate error and the data used to make the estimate (here Y) are uncorrelated.

75. The orthogonality condition gives

$$E[\epsilon X_1] = 0 \text{ and } E[\epsilon X_2] = 0.$$

Since $\epsilon = Y - \hat{Y} = Y - (\alpha_1 X_1 + \alpha_2 X_2)$, we get

$$\begin{aligned} E[\epsilon X_1] &= E[\{Y - (\alpha_1 X_1 + \alpha_2 X_2)\}X_1] \\ &= E[YX_1] - \alpha_1 E[X_1^2] - \alpha_2 E[X_1 X_2] \\ &= \rho_1\sigma_1\sigma_Y - \alpha_1\sigma_1^2 - \alpha_2\rho_{12}\sigma_1\sigma_2 \\ &= 0 \qquad\qquad \text{(given)}. \end{aligned}$$

Similarly, equating $E[\epsilon X_2] = 0$, we get $\rho_2\sigma_2\sigma_Y - \alpha_1\rho_{12}\sigma_1\sigma_2 - \alpha_2\sigma_2^2 = 0$. The two linear equations in α_1 and α_2 are given by

$$\begin{aligned} \sigma_1^2\alpha_1 + \rho_{12}\sigma_1\sigma_2\alpha_2 &= \rho_1\sigma_1\sigma_Y, \\ \rho_{12}\sigma_1\sigma_2\alpha_1 + \sigma_2^2\alpha_2 &= \rho_2\sigma_2\sigma_Y. \end{aligned}$$

In matrix notation, this can be represented as

$$\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \rho_1\sigma_1\sigma_Y \\ \rho_2\sigma_2\sigma_Y \end{pmatrix}.$$

If X_1 and X_2 are not perfectly correlated, i.e., $\rho_{12} \neq 1$, then the covariance matrix of X_1 and X_2 will be invertible. The solution to above equation is given by

$$\begin{aligned} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} &= \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1\sigma_1\sigma_Y \\ \rho_2\sigma_2\sigma_Y \end{pmatrix} \\ &= \frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)} \begin{pmatrix} \sigma_2^2 & -\rho_{12}\sigma_1\sigma_2 \\ -\rho_{12}\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \begin{pmatrix} \rho_1\sigma_1\sigma_Y \\ \rho_2\sigma_2\sigma_Y \end{pmatrix} \\ &= \frac{1}{(1-\rho_{12}^2)} \begin{pmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho_{12}}{\sigma_1\sigma_2} \\ \frac{-\rho_{12}}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{pmatrix} \begin{pmatrix} \rho_1\sigma_1\sigma_Y \\ \rho_2\sigma_2\sigma_Y \end{pmatrix} \\ &= \frac{1}{(1-\rho_{12}^2)} \begin{pmatrix} \frac{\rho_1\sigma_Y}{\sigma_1} - \rho_{12}\frac{\rho_2\sigma_Y}{\sigma_1} \\ \frac{\rho_2\sigma_Y}{\sigma_2} - \rho_{12}\frac{\rho_1\sigma_Y}{\sigma_2} \end{pmatrix}. \end{aligned}$$

Chapter 5 solutions

1. The joint density function of n variables $x_i, i = 1, \ldots, n$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = Ke^{-\mathbf{x}^T \mathbf{\Lambda}} u(\mathbf{x}) = Ke^{-\sum_{i=1}^{n} x_i \lambda_i} u(x_1) \ldots u(x_n).$$

Clearly, if K is a non-negative constant, the pdf also would be non-negative. In order that $f_{\mathbf{X}}(\mathbf{x})$ be a pdf, it should also integrate to 1. Therefore,

$$\int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} = 1 = K \int_{0}^{\infty} e^{-x_1\lambda_1}\, dx_1 \ldots \int_{0}^{\infty} e^{-x_n\lambda_n}\, dx_n = \frac{K}{\prod_{i=1}^{n} \lambda_1}.$$

Hence only for $K = \prod_{i=1}^{n} \lambda_i$ is $f_{\mathbf{X}}(\mathbf{x})$ a valid pdf.

2. Since $B_i, i = 1, \ldots, n$ are exhaustive, we can write $\Omega = \cup_{i=1}^{n} B_i$, and since they are also disjoint, $\{B_i\}_{i=1}^{n}$ form a partition. So we can write the event $\{\mathbf{X} \leq \mathbf{x}\}$ as

$$\{\mathbf{X} \leq \mathbf{x}\} = \{\mathbf{X} \leq \mathbf{x}\} \cap \Omega = \{\mathbf{X} \leq \mathbf{x}\} \cap (\cup_{i=1}^{n} B_i) = \cup_{i=1}^{n} \{\mathbf{X} \leq \mathbf{x}, B_i\}.$$

Since B_i are disjoint, $\{\mathbf{X} \leq \mathbf{x}, B_i\}$ are also disjoint, and therefore

$$\begin{aligned}
F_{\mathbf{X}}(\mathbf{x}) &= P[\{\mathbf{X} \leq \mathbf{x}\}] \\
&= P\left[\cup_{i=1}^{n} \{\mathbf{X} \leq \mathbf{x}, B_i\}\right] \\
&= \sum_{i=1}^{n} P\left[\{\mathbf{X} \leq \mathbf{x}, B_i\}\right] \\
&= \sum_{i=1}^{n} P\left[\mathbf{X} \leq \mathbf{x} | B_i\right] P[B_i] (\text{By definition of conditional probability}) \\
&= \sum_{i=1}^{n} F_{\mathbf{X}}(\mathbf{x}|B_i) P[B_i]. (\text{Definition of conditional CDF})
\end{aligned}$$

3. We note that this multi-dimensional Gaussian is factorable so that

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{x_1^2}{2\sigma_1^2}} \ldots \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{x_n^2}{2\sigma_n^2}}.$$

With $g(x, \sigma) \triangleq \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$ (the Gaussian pdf with mean 0 and variance σ^2), we write

$$f_{\mathbf{X}}(\mathbf{x}) = g(x_1, \sigma_1) \ldots g(x_n, \sigma_n),$$

where $\int_{-\infty}^{\infty} g(x_i, \sigma_1)\, dx_i = 1$. Any marginal pdf $g(x_j, \sigma_j)$ can be obtained by

$$\begin{aligned}
f_{X_j}(x_j) &= \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_n \\
&= g(x_j, \sigma_j) \int_{-\infty}^{\infty} g(x_1, \sigma_1)\, dx_1 \ldots \int_{-\infty}^{\infty} g(x_{j-1}, \sigma_{j-1})\, dx_{j-1} \\
&\quad \times \int_{-\infty}^{\infty} g(x_{j+1}, \sigma_{j+1})\, dx_{j+1} \ldots \int_{-\infty}^{\infty} g(x_n, \sigma_n)\, dx_n \\
&= g(x_j, \sigma_j).
\end{aligned}$$

4. From Equation 5.3-1 we have

$$
\begin{aligned}
f_{Y_1Y_2Y_3}(y_1, y_2, y_3) &= 3! f_{SN}(y_1) f_{SN}(y_2) f_{SN}(y_3) \\
&= \begin{cases} \frac{3!}{(2\pi)^{3/2}} e^{-\frac{1}{2}(y_1^2+y_2^2+y_3^2)}, & y_1 < y_2 < y_3, \\ 0, & \text{else.} \end{cases}
\end{aligned}
$$

5. From Problem 5.4 we have that $f_{Y_1Y_2Y_3}(y_1, y_2, y_3) = 3! f(y_1) f(y_2) f(y_3)$ for $y_1 < y_2 < y_3$ and 0 else.To get $f_{Y_1}(y_1)$ we integrate out with respect to y_2 and y_3. Thus

$$
\begin{aligned}
f_{Y_1}(y_1) &= \int\int_{y_1<y_2<y_3} f_{Y_1Y_2Y_3}(y_1, y_2, y_3) dy_2 dy_3 \\
&= 3! f(y_1) \int\int_{y_1<y_2<y_3} f(y_2) f(y_3) dy_2 dy_3 \\
&= 3! f(y_1) \int_{y_3}^{\infty} f(y_3) \left(\int_{y_1}^{y_3} f(y_2) dy_2 \right) dy_3 \\
&= 3! f(y_1) \int_{y_3}^{\infty} f(y_3) \left(F(y_3) - F(y_1)\right) dy_3 \\
&= 3! f(y_1) \int_{y_3}^{\infty} (F(y_3) - F(y_1)) dF(y_3) \\
&= 3! f(y_1) \frac{(1 - F(y_1))^2}{2} \\
&= 3 f(y_1) \left(1 - F(y_1)\right)^2 .
\end{aligned}
$$

To get $f_{Y_2}(y_2)$ we integrate out with respect to y_1 and y_3. Thus

$$
\begin{aligned}
f_{Y_2}(y_2) &= \int\int_{y_1<y_2<y_3} f_{Y_1Y_2Y_3}(y_1, y_2, y_3) dy_1 dy_3 \\
&= 3! f(y_2) \int\int_{y_1<y_2<y_3} f(y_1) f(y_3) dy_1 dy_3 \\
&= 3! f(y_2) \int_{y_2}^{\infty} f(y_3) \left(\int_{-\infty}^{y_2} f(y_1) dy_1 \right) dy_3 \\
&= 3! f(y_2) \int_{y_2}^{\infty} f(y_3) F(y_2) dy_3 \\
&= 3! f(y_2) F(y_2) \int_{y_2}^{\infty} f(y_3) dy_3 \\
&= 3! f(y_2) F(y_2) \left(1 - F(y_2)\right).
\end{aligned}
$$

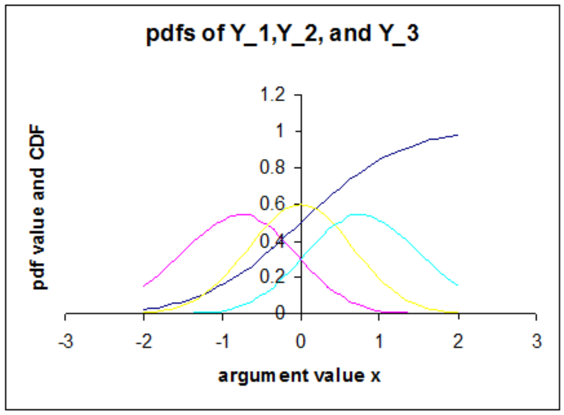


Figure 1:

Finally to get $f_{Y_3}(y_3)$, we integrate out with respect to y_1 and y_2, thus

$$
\begin{aligned}
f_{Y_3}(y_3) &= \int\int_{y_1<y_2<y_3} f_{Y_1Y_2Y_3}(y_1,y_2,y_3)dy_1dy_2 \\
&= 3!f(y_3)\int\int_{y_1<y_2<y_3} f(y_1)f(y_2)dy_1dy_2 \\
&= 3!f(y_3)\int_{-\infty}^{y_3} f(y_2)\left(\int_{-\infty}^{y_2} f(y_1)dy_1\right)dy_2 \\
&= 3!f(y_3)\int_{-\infty}^{y_3} f(y_2)F(y_2)dy_3 \\
&= 3!f(y_3)\int_{-\infty}^{y_3} F(y_2)dF(y_2) \\
&= 3f(y_3)F^2(y_3).
\end{aligned}
$$

In the figure above the pink curve is $f_{Y_1}(x)$, the yellow curve is $f_{Y_2}(x)$, and the blue curve is $f_{Y_3}(x)$. The CDF is shown in dark blue.

6. We are given $f_{Z_1Z_2\cdots Z_n}(z_1,z_2,\cdots,z_n) = n!z_1 < z_2 < \cdots < z_n$ and 0 else. To get $f_{Z_1Z_n}(z_1,z_n)$ we integrate out with respect to $z_2, z_3, ..., z_{n-1}$. Thus

$$f_{Z_1Z_n}(z_1,z_n) = n!\int_{z_{n-1}}^{z_n}\left(\cdots\left(\int_{z_1}^{z_3} dz_2\right)\cdots\right)dz_{n-1}.$$

Let's take the first integration and leave out the $n!$:$\int_{z_1}^{z_3} dz_2 = z_3 - z_1$. The second integration yields $\int_{z_1}^{z_4}(z_3-z_1)dz_3 = \frac{(z_4-z_1)^2}{2}$. The third integration yields $\int_{z_1}^{z_5}(z_4-z_1)^2/2dz_3 = \frac{(z_5-z_1)^3}{3\cdot 2}$. Thus after $n-2$ integrations we end up with $\int_{z_1}^{z_n}\frac{(z_{n-1}-z_1)^{n-3}}{n-3}dz_{n-1} = \frac{(z_n-z_1)^{n-2}}{(n-2)\cdots 3\cdot 2}$. Putting back the $n!$ then yields

$$f_{Z_1Z_n}(z_1,z_n) = \begin{cases} n(n-1)(z_n-z_1)^{n-2}, & 0<z_1<z_n<1, n\geq 2 \\ 0, & \text{else.} \end{cases}$$

7. Let $V_{1n} \triangleq Z_n - Z_1$ and define the auxiliary random variable $W \triangleq Z_1$. Consider the functional equations $v = z_n - z_1, w = z_1$. The Jacobian of this transformation is $\begin{vmatrix} \partial v/\partial z_1 & \partial v/\partial z_n \\ \partial w/\partial z_1 & \partial w/\partial z_n \end{vmatrix} = \begin{vmatrix} -1 & 1 \\ 0 & 1 \end{vmatrix}$, so that $|J| = 1$. Substituting into $f_{Z_1Z_n}(z_1, z_n) = n(n-1)(z_n - z_1)^{n-2}$, we get

$$f_{V_{1n}W}(v, w) = \begin{cases} n(n-1)v^{n-2}, & 0 < w < 1 - v,\ 0 < v < 1, \\ 0, & \text{else.} \end{cases}$$

8. We already know that $f_{V_{1n}W}(v, w) = n(n-1)v^{n-2}, 0 < w < 1 - v, 0 < v < 1$. To get $f_{V_{1n}}(v)$, we integrate out with respect to w. This yields

$$\begin{aligned} f_{V_{1n}}(v) &= n(n-1)v^{n-2} \int_{w=0}^{1-v} dw \\ &= \begin{cases} n(n-1)(1-v)v^{n-2}, & 0 < v < 1, n \geq 2, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

9. We need to show that the area under

$$f_{Z_1Z_2Z_3}(z_1, z_2, z_3) = \begin{cases} 3!, & 0 < z_1 < z_2 < z_3 < 1, \\ 0, & \text{else,} \end{cases}$$

is 1. A simple repeated integration yields

$$\begin{aligned} 6\int_0^1 \left(\int_0^{z_3} \left(\int_0^{z_2} dz_1 \right) dz_2 \right) dz_3 &= 6\int_0^1 \left(\int_0^{z_3} z_2 dz_2 \right) dz_3 \\ &= 6(\frac{1}{2}) \int_0^1 z_3^2 dz_3 \\ &= 6(\frac{1}{2})(\frac{1}{3}) \left(z_3^3 \big|_0^1 \right) \\ &= 1. \end{aligned}$$

10. The beta pdf is given

$$f(x; \alpha, \beta) = \begin{cases} \frac{(\alpha+\beta+1)}{\alpha!\beta!} x^\alpha (1-x)^\beta, & 0 < x < 1, \\ 0, & \text{else.} \end{cases}$$

So for $\beta = 0$ and $n = \alpha + 2 = 2$, which implies that $\alpha = 0$, we get $f(x; 0, 0) = 1, 0 < x < 1$. Hence the CDF is

$$F(x; 0, 0) = \begin{cases} 0, x \leq 0 \\ x, 0 < x \leq 1 \\ 1, x > 1 \end{cases} .$$

11. First we derive Equation 5.3-11. We begin with $f_{Z_1Z_2Z_3}(z_1z_2z_3) = 3!$ for $0 < z_1 < z_2 < z_3 < 1$ and 0, else. As shown in Section 5.3, $f_{Z_2Z_3}(z_2z_3) = 3!z_2$ for $0 < z_2 < z_3 < 1$. Now let $V_{23} \triangleq Z_3 - Z_2$ and $W \triangleq Z_2$, with the functional equations $v = z_2 - z_3$ and $w = z_2$. We find $\partial v/\partial z_2 = -1; \partial v/\partial z_3 = 1; \partial w/\partial z_2 = 1; \partial w/\partial z_3 = 0$. Hence $|J| = 1$ and

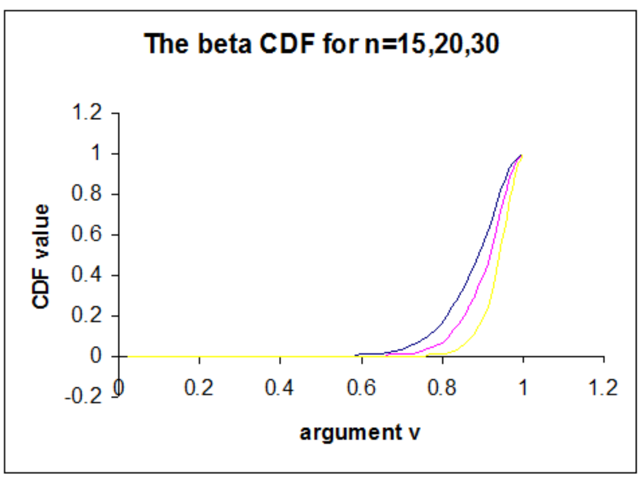
$$f_{V_{23}W}(v, w) = \begin{cases} 3!w, & 0 < w < 1 - v, 0 < v < 1 \\ 0, & \text{else.} \end{cases} .$$

Finally

$$\begin{aligned} f_{V_{23}}(v) &= \int_0^{1-v} 3!wdw \\ &= \begin{cases} 3!(1-v)^2/2, & 0<v<1, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

Next we derive Equation 5.3-12. We compute $f_{Z_1Z_2}(z_1,z_2) = 3!\int_{z_2}^{1} dz_3 = 3!(1-z_2)$ for $0<z_1<z_2<1$. Let $V_{12} \triangleq Z_2 - Z_1$ and $W \triangleq Z_1$, with the functional equations $v = z_2 - z_1$ and $w = z_1$. Once again we find that $|J| = 1$ and $f_{V_{12}W}(v,w) = 3!(1-w-v)$ for $0<w<1-v, 0<v<1$. Integrating out with respect to w yields $3!\int_0^{1-v} wdw = 3!(1-v)^2/2$, the same as before! To derive Equation 5.3-13 takes more work but follows the same procedure as in deriving Equations 5.3-11 and 5.3-12. We let $V_{lm} \triangleq Z_m - Z_l$ and $W \triangleq Z_l$, with functional equations $v = z_m - z_l$ and $w = z_l$. The magnitude of the Jacobian of this transformation is 1. To get $f_{V_{lm}W}(v,w)$ we have to integrate out with respect to $z_1, z_2, ..., z_{l-1}$; then integrate out with respect to $z_{l+1}, ..., z_{m-1}$; and finally integrate out with respect to $z_{m+1}, ..., z_n$. When we are done with these integrations we integrate out with respect to w to obtain Equation 5.3-13.

12. The beta CDF for $\beta = 1$ *and* $n =$ is shown in the figure below. Recall that $V_{1n} \triangleq \int_{Y_1}^{Y_n} f_X(x)dx$ and therefore V_{1n} is the probability area between the smallest and largest of the ordered values $Y_1, Y_2, \cdots, Y_n$ obtained from the unordered i.i.d. observations on X.The beta CDF is $F_{V_{1n}}(v)=P[\int_{Y_1}^{Y_n} f_X(x)dx \le v]$. We note that for large $nF_{V_{1n}}(v)$is exceedingly small except for argument values approaching 1. Why is that? For large n it is exceedingly unlikely that the smallest and largest value of the observations on Xwill be near each other. If they were the probability area bounded by them would be very small. It is much more likely that for large n Y_1 will be very small and $Y_n$will be very large. This would cause the probability area between them to be near 1. In effect this is what the curves in the figure below show.



13. There are several ways to do this problem. Here is one: Write

$$E[Z_1] = n!\int_0^1\int_0^{z_n}\cdots\int_0^{z_2} z_1 dz_1\cdots dz_{n-1}dz_n$$

Now, let's do the integrations one-by-one, leaving out temporarily the factor $n!$

$$\begin{aligned}
\int_0^{z_2} z_1 dz_1 &= \frac{z_2^2}{2} \\
\int_0^{z_3} (z_2^2/2) dz_1 &= \frac{z_3^3}{3 \cdot 2} \\
&\vdots \\
\int_0^{z_n} \frac{z_{n-1}^{n-1}}{(n-1)!} dz_{n-1} &= \frac{z_n^n}{n!}, \\
\text{and finally, } \int_0^1 \frac{z_n^n}{n!} dz_n &= \frac{1}{(n+1)!}
\end{aligned}$$

Hence

$$\begin{aligned}
E[Z_1] &= n! \int_0^1 \int_0^{z_n} \cdots \int_0^{z_2} z_1 dz_1 \cdots dz_{n-1} dz_n \\
&= \frac{n!}{(n+1)!} = \frac{1}{n+1}.
\end{aligned}$$

In a similar way we could compute

$$\begin{aligned}
E[Z_2] &= \frac{2}{n+1} \\
E[Z_3] &= \frac{3}{n+1} \\
&\vdots \\
\text{and } E[Z_n] &= \frac{n}{n+1}.
\end{aligned}$$

Here is a better way: Write

$$\begin{aligned}
Z_i &= (Z_i - Z_1) + Z_1 \\
E[Z_i] &= E[Z_i - Z_1] + E[Z_1] \\
&= E[V_{1i}] + \frac{1}{n+1} \\
&= \frac{n!}{(i-2)!(n-i+1)!} \int_0^1 v \cdot \left(v^{i-2} (1-v)^{n-i+1} \right) dv + \frac{1}{n+1},
\end{aligned}$$

by Equation5.3-13. The integral $\int_0^1 v^{i-1}(1-v)^{n-i+1} dv$ is evaluated in any table of definite available on the internet, for example, formula 7 at sosmath.com/tables/integral/integ41/integ41/html as

$$\int_0^1 \left(v^{i-1} (1-v)^{n-i+1} \right) dv = \frac{(i-1)!(n-i-1)!}{(n+1)!}$$

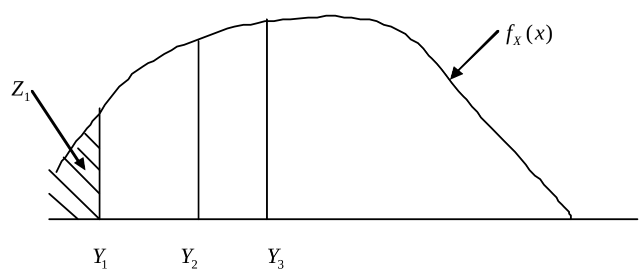
Hence

$$\begin{aligned}
E[Z_i] &= \frac{i-1}{n+1} + \frac{1}{n+1} \\
&= \frac{i}{n+1}.
\end{aligned}$$

14. Consider the n ordered RVs $Y_1 < Y_2 < \cdots < Y_n$ ordered from $X_1, X_2, \cdots, X_n$. With $Z_i = \int_{-\infty}^{Y_i} f_X(x)dx$, we found in the previous problem that $E[Z_i] = \frac{i}{n+1}$. Hence the average probability area between adjacent parts is

$$\begin{aligned} E[Z_{i+1}] - E[Z_i] &= \frac{i+1}{n+1} - \frac{i}{n+1} \\ &= \frac{1}{n+1}. \end{aligned}$$

Thus the number of equal parts is $1/\left(\frac{1}{n+1}\right) = n+1$.



15. (a) Let $\mathbf{x} = (x_1, x_2, ..., x_n)^T$, then

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} &= \begin{bmatrix} x_1x_1 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & x_2x_2 & \cdots & x_2x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_nx_1 & x_nx_2 & & x_nx_n \end{bmatrix} \\ &= \begin{bmatrix} x_1\mathbf{x}^T \\ x_2\mathbf{x}^T \\ \vdots \\ x_n\mathbf{x}^T \end{bmatrix}. \end{aligned}$$

Hence any row is obtained from any other row by a scalar multiplication. Therefore, there is at most one linearly independent row. Thus the rank is at most 1.

(b) The expectation operator tends to destroy the linear dependence, i.e.

$$E[\mathbf{X}\mathbf{X}^T] = \begin{bmatrix} E[X_1X_1] & \cdots & E[X_1X_n] \\ \vdots & \cdots & \vdots \\ E[X_nX_1] & \cdots & E[X_nX_n] \end{bmatrix},$$

and we cannot usually write any row as a linear combination of any other one. Indeed if

$$E[X_iX_j] \simeq \frac{1}{N}\sum_{k=1}^{N} x_i^{(k)} x_j^{(k)},$$

when N is large, the matrix can be full rank, provided that $N > n$.

16. We need to show that the CDF of $Y_n \triangleq \max(X_1, X_2, ..., X_n)$ is $F_{Y_n}(y) = F_X^n(y)$. Here the $X_1, X_2, ..., X_n$ are n i.i.d. observations on X with CDF $F_X(x)$. Clearly $F_{Y_n}(y) = P[\max(X_1, X_2, ..., X_n) \leq y]$. But if the max is less than y, all the other RVs must be less than y, hence

$$\begin{aligned} F_{Y_n}(y) &= P[\max(X_1, X_2, ..., X_n) \leq y] \\ &= P[X_1 \leq y]P[X_2 \leq y] \cdots P[X_n \leq y] \\ &= F_X^n(y). \end{aligned}$$

Note that in obtaining the last line we used that the Xs were i.i.d. with CDF $F_X(x)$.

17. Recalling that $Y_1 \triangleq \min(X_1, X_2, ..., X_n)$, we note that $P[Y_1 \leq y] = 1 - P[Y_1 > y]$. Now if the *smallest* i.e. Y_1, of the Xs is *greater* than y, then all the other Xs must be greater than y also. Hence

$$\begin{aligned} P[Y_1 &> y] = P[X_1 > y, X_2 > y ... X_n > y] \\ &= P[X_1 > y]P[X_2 > y] \cdots P[X_n > y] \\ &= (1 - F_{X_1}(y))(1 - F_{X_2}(y)) \cdots (1 - F_{X_n}(y)) \\ &= (1 - F_X(y))^n . \end{aligned}$$

In obtaining the last line, we used the fact that the $\{X_i, i = 1, ..., n\}$ are i.i.d. with CDF $F_X(x)$. Finally, because $P[Y_1 \leq y] = 1 - P[Y_1 > y]$, there results that

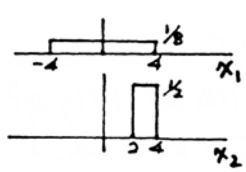
$$\begin{aligned} P[Y_1 \leq y] &= 1 - (1 - F_X(y))^n \\ &\triangleq F_{Y_1}(y). \end{aligned}$$

18. Given the ordered samples $Y_1, Y_2, ..., Y_r, ..., Y_n$ from the n i.i.d. observations $X_1, X_2, ..., X_n$ on X, we want to compute $P[Y_r \leq y]$. However, the event $\{Y_r \leq y\}$ implies that *at least* r of the unordered $\{X_i\}$ satisfy $\{X_i \leq y\}$. Now $P[X_i \leq y]$ is merely $F_X(y)$, and if we count it as a "success" when$\{X_i \leq y\}$, then the probability of at least r successes is the binomial probability

$$P[Y_r \leq y] = \sum\nolimits_{i=r}^{n} \begin{pmatrix} n \\ i \end{pmatrix} F_X^i(y)\,(1 - F_X(y))^{n-1}.$$

19. We can write

$$\begin{aligned} f_{X_1}(x_1) &= \frac{1}{8}[u(x_1 + 4) - u(x_1 - 4)] \\ f_{X_2}(x_2) &= \frac{1}{8}[u(x_2 - 2) - u(x_2 - 4)]. \end{aligned}$$



Then $f_{X_1}(x_1)f_{X_2}(x_2) = f_{X_1,X_2}(x_1,x_2)$ and the RVs X_1 and X_2 are independent. Hence, they are uncorrelated,

$$\begin{aligned} E[X_1X_2] &= E[X_1]E[X_2] \\ &= 0 \end{aligned}$$

since $E[X_1] = 0$. Thus, X_1 and X_2 are orthogonal.

20. Since $\mathbf{X}_i$ and $\mathbf{X}_j$ are mutually orthogonal, $E[\mathbf{X}_i^T\mathbf{X}_j] = 0$ for any $i \neq j, 1 \leq i, j \leq n$.

$$\begin{aligned} E\left[\left\|\sum_{i=1}^{n}\mathbf{X}_i\right\|^2\right] &= E\left[(\mathbf{X}_1 + \ldots + \mathbf{X}_n)^T(\mathbf{X}_1 + \ldots + \mathbf{X}_n)\right] \\ &= E\left[\mathbf{X}_1^T\mathbf{X}_1 + \ldots + \mathbf{X}_n^T\mathbf{X}_n + \sum_{i=1}^{n}\sum_{j\neq i}^{n}\mathbf{X}_i^T\mathbf{X}_j\right] \\ &= E\left[\mathbf{X}_1^T\mathbf{X}_1 + \ldots + \mathbf{X}_n^T\mathbf{X}_n\right] + E\left[\sum_{i=1}^{n}\sum_{j\neq i}^{n}\mathbf{X}_i^T\mathbf{X}_j\right] \\ &= \sum_{i=1}^{n}E[\mathbf{X}_i^T\mathbf{X}_i] + \sum_{i=1}^{n}\sum_{j\neq i}^{n}E[\mathbf{X}_i^T\mathbf{X}_j] \\ &= \sum_{i=1}^{n}E[\mathbf{X}_i^T\mathbf{X}_i] + 0 \\ &= \sum_{i=1}^{n}E[\mathbf{X}_i^T\mathbf{X}_i] \\ &= \sum_{i=1}^{n}E[\|\mathbf{X}_i\|]. \end{aligned}$$

21. Since the random vectors are mutually uncorrelated, $E[(\mathbf{X}_i - \boldsymbol{\mu}_i)(\mathbf{X}_j - \boldsymbol{\mu}_j)^T] = \mathbf{0}$ for all $i \neq j, 1 \leq i, j \leq n$.

$$\begin{aligned} E\left[\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu}_i)\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{\mu}_j)^T\right] &= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu}_i)(\mathbf{X}_j - \boldsymbol{\mu}_j)^T\right] \\ &= E\left[\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu}_i)(\mathbf{X}_i - \boldsymbol{\mu}_i)^T + \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}(\mathbf{X}_i - \boldsymbol{\mu}_i)(\mathbf{X}_j - \boldsymbol{\mu}_j)\right] \\ &= \sum_{i=1}^{n}\mathbf{K}_{ii} + E\left[\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}(\mathbf{X}_i - \boldsymbol{\mu}_i)(\mathbf{X}_j - \boldsymbol{\mu}_j)\right] \\ &= \sum_{i=1}^{n}\mathbf{K}_{ii} + \mathbf{0}. \end{aligned}$$

22.

$$\begin{aligned}\sum_{i=1}^{n}(\mathbf{X}_i-\boldsymbol{\mu}_i)\sum_{j=1}^{n}(\mathbf{X}_j-\boldsymbol{\mu}_j)^T &= \sum_{i=1}^{n}\sum_{j=1}^{n}(\mathbf{X}_i-\boldsymbol{\mu}_i)(\mathbf{X}_j-\boldsymbol{\mu}_j)^T \\ &= \sum_{i=1}^{n}(\mathbf{X}_i-\boldsymbol{\mu}_i)(\mathbf{X}_i-\boldsymbol{\mu}_i)^T+\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}(\mathbf{X}_i-\boldsymbol{\mu}_i)(\mathbf{X}_j-\boldsymbol{\mu}_j).\end{aligned}$$

So, upon taking expectations, we have

$$\begin{aligned}E\left[\sum_{i=1}^{n}(\mathbf{X}_i-\boldsymbol{\mu}_i)\sum_{j=1}^{n}(\mathbf{X}_j-\boldsymbol{\mu}_j)^T\right] &= \sum_{i=1}^{n}\mathbf{K}_i+\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}E[(\mathbf{X}_i-\boldsymbol{\mu}_i)(\mathbf{X}_j-\boldsymbol{\mu}_j)] \\ &= \sum_{i=1}^{n}\mathbf{K}_i+\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}E[(\mathbf{X}_i-\boldsymbol{\mu}_i)]E[(\mathbf{X}_j-\boldsymbol{\mu}_j)] \\ &= \sum_{i=1}^{n}\mathbf{K}_i+\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}(E[\mathbf{X}_i]-\boldsymbol{\mu}_i)](E[\mathbf{X}_j]-\boldsymbol{\mu}_j) \\ &= \sum_{i=1}^{n}\mathbf{K}_i+0.\end{aligned}$$

23. (a) From the Schwarz inequality,

$$\begin{aligned}\sigma_i^2\sigma_j^2 &= E[(X_i-\mu_i)^2]E[(X_j-\mu_j)^2] \\ &\geq \left|E[(X_i-\mu_i)(X_j-\mu_j)]\right|^2 \\ &= |K_{ij}|^2.\end{aligned}$$

But in the given matrix, $\sigma_1^2\sigma_2^2 = 2\times 3 = 6$ and $|K_{12}|^2 = 16 > \sigma_1^2\sigma_2^2$. Thus violating the Schwarz inequality.

(b) Always $\sigma_{33}^2 = E[(X_3-\mu_3)^2] \geq 0$. In the given matrix, this value is -2.

(c) The covariance value $K_{12} = E[(X_1-\mu_1)(X_2-\mu_2)]$ must be real for a real valued random vector. In the given matrix, there are numbers with non-zero imaginary parts.

(d) Always for covariance matrices of real valued random vectors, we have the symmetry conditions

$$\begin{aligned}K_{ij} &= E[(X_i-\mu_i)(X_j-\mu_j)] \\ &= E[(X_j-\mu_j)(X_i-\mu_i)] \\ &= K_{ji}.\end{aligned}$$

But, in the given matrix, $K_{23} = 3, K_{32} = 12 \neq K_{23}$, thus violating the required symmetry.

24. (a)

$$\begin{aligned}\det(\mathbf{K_{XX}}-\lambda\mathbf{I}) &= \det\begin{bmatrix}3-\lambda & \sqrt{2}\\ \sqrt{2} & 4-\lambda\end{bmatrix} = 0 \\ \Rightarrow &\quad \lambda^2-7\lambda+10=0,\ \text{or}\ \ \lambda_1=5,\ \ \text{and}\ \ \lambda_2=2.\end{aligned}$$

i) $\lambda_1 = 5:$ $\quad \mathbf{K_{XX}}\boldsymbol{\phi}_1 = 5\boldsymbol{\phi}_1$ or what is the same, $(\mathbf{K_{XX}} - 5\mathbf{I}) = \mathbf{0}$ leads to

$$\begin{aligned} -2\phi_{11} + \sqrt{2}\phi_{12} = & \quad 0 \\ \sqrt{2}\phi_{11} - \phi_{12} = & \quad 0 \\ \Rightarrow \quad \boldsymbol{\phi}_1 & = (\phi_{11}, \phi_{12})^T \\ & = \frac{1}{\sqrt{3}}(1, \sqrt{2})^T \quad . \end{aligned}$$

ii) $\lambda_2 = 2:$ $\quad \mathbf{K_{XX}}\boldsymbol{\phi}_1 = 2\boldsymbol{\phi}_1$ or what is the same, $(\mathbf{K_{XX}} - 2\mathbf{I}) = \mathbf{0}$ leads to

$$\begin{aligned} \phi_{21} + \sqrt{2}\phi_{22} = & \quad 0 \\ \sqrt{2}\phi_{21} + 2\phi_{22} = & \quad 0 \\ \Rightarrow \quad \boldsymbol{\phi}_2 & = (\phi_{21}, \phi_{22})^T \\ & = \sqrt{\frac{2}{3}}(1, -\frac{1}{\sqrt{2}})^T \quad . \end{aligned}$$

Thus

$$\begin{aligned} \mathbf{\Phi} &= \begin{bmatrix} \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{3}} \end{bmatrix}. \end{aligned}$$

Set

$$\begin{aligned} \mathbf{\Lambda}^{-\frac{1}{2}} &\triangleq \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, \end{aligned}$$

and the define

$$\mathbf{C} \triangleq \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{\Phi}^T.$$

Then with $\mathbf{Y} = \mathbf{CX}$, we have

$$\begin{aligned} E[\mathbf{YY}^T] &= \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{\Phi}^T E[\mathbf{XX}^T]\mathbf{\Phi\Lambda}^{-\frac{1}{2}} \\ &= \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{\Phi}^T\mathbf{K_{XX}}\mathbf{\Phi\Lambda}^{-\frac{1}{2}} \\ &= \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{\Phi}^T\left(\mathbf{\Phi\Lambda}\right)\mathbf{\Lambda}^{-\frac{1}{2}} \\ &= \mathbf{\Lambda}^{-\frac{1}{2}}(\mathbf{\Phi}^T\mathbf{\Phi})\mathbf{\Lambda\Lambda}^{-\frac{1}{2}} \\ &= \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{I\Lambda\Lambda}^{-\frac{1}{2}} \\ &= \mathbf{I}. \end{aligned}$$

Thus $\mathbf{Y} = \mathbf{CX} \;\Rightarrow\; \mathbf{K_{XX}} = \mathbf{I}$. Finally, since $\mathbf{C}^T \neq \mathbf{C}^{-1}$, $\mathbf{C}$ is not unitary.

(b) i)

$$\mathbf{AA}' = \begin{bmatrix} aa' + bb' & ab' + bc' \\ a'b + b'c & bb' + cc' \end{bmatrix} \quad \text{and} \quad \mathbf{A}'\mathbf{A} = \begin{bmatrix} aa' + bb' & a'b + b'c \\ ab' + c'b & bb' + cc' \end{bmatrix}.$$

Now

$$\begin{aligned} ab' + bc' &= a'b + b'c \quad \text{if} \quad a = c,\ a\prime = c\prime \\ a'b + b'c &= ab' + c'b \quad \text{if} \quad a = c,\ a\prime = c. \end{aligned}$$

ii) $(\mathbf{AA'})^T = \mathbf{A'}^T\mathbf{A}^T = \mathbf{A'A}$ since $\mathbf{A'}^T = \mathbf{A'}, \mathbf{A} = \mathbf{A}^T$. Hence if $\mathbf{A'A} = \mathbf{AA'}$, then $(\mathbf{AA'})^T = \mathbf{AA'}$, and the product is a real symmetric matrix.

25. (a) We know $\mathbf{KA} = (\mathbf{A}^T)^{-1}$, and so

$$\begin{aligned} \mathbf{K}_1\mathbf{A} &= (\mathbf{A}^T)^{-1}\mathbf{\Lambda}^{(1)} \\ &= \mathbf{KA\Lambda}^{(1)}. \end{aligned}$$

Thus

$$\mathbf{K}^{-1}\mathbf{K}_1\mathbf{A} = \mathbf{A\Lambda}^{(1)}.$$

(b) Now

$$\begin{aligned} \mathbf{A}^T\mathbf{KA} &= \mathbf{I} \\ &= a_1\underbrace{\mathbf{A}^T\mathbf{K}_1\mathbf{A}}_{\mathbf{\Lambda}^{(1)}} + a_2\mathbf{A}^T\mathbf{K}_2\mathbf{A} \end{aligned}$$

So

$$\begin{aligned} \mathbf{A}^T\mathbf{K}_2\mathbf{A} &= \frac{1}{a_2}[\mathbf{I} - a_1\mathbf{\Lambda}^{(1)}], \quad \text{a difference of diagonal matrices,} \\ &= \mathbf{\Lambda}^{(2)}, \quad \text{a diagonal matrix.} \end{aligned}$$

(c) All diagonal matrices of the same size share the same eigenvectors:

$$\begin{aligned} \boldsymbol{\phi}_1 &= (1, 0, 0, \ldots, 0)^T, \\ \boldsymbol{\phi}_2 &= (0, 1, 0, \ldots, 0)^T, \\ &\ \ \vdots \qquad\qquad\qquad , \\ \boldsymbol{\phi}_n &= (0, \ldots, 0, 0, 1)^T. \end{aligned}$$

(d)

$$\mathbf{\Lambda}^{(2)} = \frac{1}{a_2}[\mathbf{I} - a_1\mathbf{\Lambda}^{(1)}] \quad \text{or} \quad \lambda_i^{(2)} = \frac{1}{a_2}(1 - a_1\lambda_i^{(1)}),$$

so $\max \lambda_i^{(1)}$ produces $\min \lambda_i^{(2)}$ if both $a_1 > 0$ and $a_2 > 0$.

26. (a) Write $\mathbf{W} = (\mathbf{X}_1 : \mathbf{X}_2 : \cdots : \mathbf{X}_m)$. Then

$$\begin{aligned} \mathbf{WW}^T &= [\mathbf{X}_1 : \mathbf{X}_2 : \cdots : \mathbf{X}_m] \begin{bmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \\ \vdots \\ \mathbf{X}_m^T \end{bmatrix} \\ &= \left[\mathbf{X}_1\mathbf{X}_1^T + \mathbf{X}_2\mathbf{X}_2^T + \ldots + \mathbf{X}_m\mathbf{X}_m^T\right]. \end{aligned}$$

Hence

$$\frac{1}{m}\mathbf{WW}^T = \frac{1}{m}\sum_{i=1}^{m}\mathbf{X}_i\mathbf{X}_i^T.$$

(b) Consider the case when $m = 2, n = 3$. Then with

$$\mathbf{X}_1 \triangleq (X_{11}, X_{21}, X_{31})^T \quad \text{and} \quad \mathbf{X}_2 \triangleq (X_{12}, X_{22}, X_{32})^T,$$

we get

$$[\mathbf{X}_1 : \mathbf{X}_2] \begin{bmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \end{bmatrix} = \begin{bmatrix} X_{11}\mathbf{X}_1^T + X_{12}\mathbf{X}_2^T \\ X_{21}\mathbf{X}_1^T + X_{22}\mathbf{X}_2^T \\ X_{31}\mathbf{X}_1^T + X_{32}\mathbf{X}_2^T \end{bmatrix}.$$

Thus there are three rows, but only two independent row vectors. So, any row can be written as a linear combination of the other two rows. Thus the rank is 2. More generally, we will have

$$\mathbf{S} = \begin{bmatrix} X_{11}\mathbf{X}_1^T + \ldots + X_{1m}\mathbf{X}_m^T \\ \vdots \\ X_{n1}\mathbf{X}_1^T + \ldots + X_{nm}\mathbf{X}_m^T \end{bmatrix}, \text{ with } \operatorname{rank}(\mathbf{S}) \leq m.$$

(c) We have $\mathbf{W} = (n \times m)$ and so $\mathbf{W}^T = (m \times n)$. Hence

$$\mathbf{S}' = \frac{1}{m}\mathbf{W}^T\mathbf{W} = (m \times n).$$

Write

$$\begin{aligned} \mathbf{S\Phi} &= \mathbf{\Phi\Lambda} = \frac{1}{m}\mathbf{W}\mathbf{W}^T\mathbf{\Phi} \\ \mathbf{S}'\mathbf{\Phi}' &= \frac{1}{m}\mathbf{W}^T\mathbf{W}\mathbf{\Phi}' = \mathbf{\Phi}'\mathbf{\Lambda}'. \end{aligned}$$

Then

$$\begin{aligned} \mathbf{W}\mathbf{S}'\mathbf{\Phi}' &= \mathbf{W}\mathbf{\Phi}'\mathbf{\Lambda}' \\ &= \frac{1}{m}\mathbf{W}\mathbf{W}^T(\mathbf{W}\mathbf{\Phi}') \\ &= (\mathbf{W}\mathbf{\Phi}')\mathbf{\Lambda}'. \end{aligned}$$

Thus, the first m eigenvalues of $\mathbf{S}\prime$ are those of $\mathbf{S}$, and the remaining $n - m$ eigenvalues are zero. The first m eigenvalues of $\mathbf{S}\prime$ are $\mathbf{W}\mathbf{\Phi}' = (n \times m) \times (m \times n) = (n \times n)$.

(d) Only an $m \times m$ matrix is used for calculating m eigenvalues. In practice, very often $n >> m$, so this method saves a lot of work.

27.

28.

29. The mean of $\mathbf{Y}$ is given by

$$E[\mathbf{Y}] = E[\mathbf{A}^T\mathbf{X} + B] = \mathbf{A}^T\boldsymbol{\mu} + B = \begin{pmatrix} 2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ -5 \\ 6 \end{pmatrix} + 5 = 32.$$

Let $E[\mathbf{Y}] = \boldsymbol{\mu}_\mathbf{1}$. Then

$$\begin{aligned}
\sigma_Y^2 &= E\left[(\mathbf{Y}-\boldsymbol{\mu}_\mathbf{1})^T(\mathbf{Y}-\boldsymbol{\mu}_\mathbf{1})\right] \\
&= E\left[\left(\mathbf{A}^T(\mathbf{X}-\boldsymbol{\mu})\right)^T\left(\mathbf{A}^T(\mathbf{X}-\boldsymbol{\mu})\right)\right] \\
&= E\left[\left((\mathbf{X}-\boldsymbol{\mu})^T\mathbf{A}\right)\left(\mathbf{A}^T(\mathbf{X}-\boldsymbol{\mu})\right)\right] \\
&= E\left[\left(\mathbf{A}^T(\mathbf{X}-\boldsymbol{\mu})\right)\left((\mathbf{X}-\boldsymbol{\mu})^T\mathbf{A}\right)\right] \\
&= \mathbf{A}^T E\left[(\mathbf{X}-\boldsymbol{\mu})^T(\mathbf{X}-\boldsymbol{\mu})\right]\mathbf{A} \\
&= \mathbf{A}^T\mathbf{K}_\mathbf{X}\mathbf{A} \\
&= \begin{pmatrix} 2 & -1 & 2 \end{pmatrix}\begin{pmatrix} 5 & 2 & -1 \\ 2 & 5 & 0 \\ -1 & 0 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} = 25.
\end{aligned}$$

30.

31. The mean of Y is given as

$$E[Y] = E[\sum_{i=1}^n p_i X_i] = \sum_{i=1}^n p_i E[X_i] = \sum_{i=1}^n p_i \mu_i.$$

The variance of Y is given by

$$\begin{aligned}
\sigma_Y^2 &= E[(Y-\mu_Y)^2] \\
&= E\left[\sum_{i=1}^n p_i(X_i-\mu_i)\right]^2 \\
&= \sum_{i=1}^n\sum_{j=1}^n p_i p_j E[(X_i-\mu_i)(X_j-\mu_j)] \\
&= \sum_{i=1}^n\sum_{j=1}^n p_i p_j K_{ij}.
\end{aligned}$$

32.

33. Let r and p be the parameters of the binomial distribution, then for each X_i we have

$$\begin{aligned}
\Phi_{X_i}(\omega) &= \sum_{k=0}^r \binom{r}{k} p^k q^{r-k} e^{+j\omega k} \\
&= \left(pe^{j\omega}+q\right)^r \quad \text{with} \quad q \triangleq 1-p.
\end{aligned}$$

Considering the vector $\mathbf{X} = (X_1, \ldots, X_n)^T$, we have

$$\begin{aligned}
\Phi_\mathbf{X}(\boldsymbol{\omega}) &= E[e^{+j\boldsymbol{\omega}^T\mathbf{X}}] \\
&= \prod_{i=1}^n E[e^{+j\omega_i X_i}] \\
&= \prod_{i=1}^n \left(pe^{j\omega_i}+q\right)^r, \qquad \text{by the i.i.d. assumption.}
\end{aligned}$$

Then, with $Y = \sum_{i=1}^{n} X_i = \mathbf{1}^T\mathbf{X}$, where is a column vector of 1's, we have

$$\begin{aligned}\Phi_Y(\omega) &= E[e^{+j\omega Y}] \\ &= E[e^{+j\omega \mathbf{1}^T\mathbf{X}}] \\ &= \Phi_{\mathbf{X}}(\omega\mathbf{1}) \\ &= \left(pe^{j\omega}+q\right)^{nr},\end{aligned}$$

which we can recognize as binomial with parameters nr and p. Hence

$$\begin{aligned}P_Y(k) &= P[Y=k] \\ &= \binom{nr}{k} p^k q^{nr-k} \qquad k=0,1,...,nr.\end{aligned}$$

The mean of Y is nr, and the variance of Y is $nrpq$.

34.

35. The first order pdf f_{X_1} is computed from

$$\begin{aligned}f_{X_1}(x) &= \frac{2}{3}\int_0^1\int_0^1 (x_1+x_2+x_3)dx_2dx_3 \\ &= \begin{cases} \frac{2}{3}(x+1), & 0<x\le 1, \\ 0, & \text{else.}\end{cases}\end{aligned}$$

By symmetry, $f_{X_1} = f_{X_2} = f_{X_3}$. The second-order pdf is calculated as

$$\begin{aligned}f_{X_1,X_2}(x,y) &= \int_0^1 (x+y+x_3)dx_3 \\ &= \begin{cases} \frac{2}{3}(x+y+\frac{1}{2}), & 0<x,y\le 1, \\ 0, & \text{else.}\end{cases}\end{aligned}$$

By symmetry, $f_{X_1,X_2} = f_{X_1,X_3} = f_{X_2,X_2}$. Likewise for the first moment, we get

$$\begin{aligned}E[X_1] &= \int_0^1 x\frac{2}{3}(x+1)dx \\ &= \frac{5}{9} = \mu_1 \\ &= \mu_2 = \mu_3.\end{aligned}$$

Then the variance is given as

$$\begin{aligned}\mathrm{Var}[X_i] &= E[X_i^2] - \mu_i^2 \\ &= \frac{2}{3}\int_0^1 x^2(x+1)dx - (\frac{5}{9})^2 \\ &\doteq 0.08 = \sigma^2.\end{aligned}$$

Also, $K_{12} = K_{13} = K_{23} \doteq -0.003$. Then, upon setting $\rho_{12} = K_{12}/\sigma^2 = \rho_{13} = \rho_{23}$, we have

$$\mathbf{K} \doteq 0.08 \begin{bmatrix} 1 & -0.04 & -0.04 \\ -0.04 & 1 & -0.04 \\ -0.04 & -0.04 & 1 \end{bmatrix}.$$

Since $\rho_{12} = \rho_{13} = \rho_{23} \simeq 0$, the random variables X_1, X_2, and X_3 are almost uncorrelated.

36.

37. Let $g_k(\mathbf{x}) \triangleq \sum_{j=1}^{n} a_{kj}x_j, j = 1, ..., n$. Then the only solution to the set of equations

$$\begin{bmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_n(\mathbf{x}) \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \mathbf{0},$$

is

$$\begin{aligned} \mathbf{x}^o &= \mathbf{A}^{-1}\mathbf{y} \\ &= \mathbf{B}\mathbf{y}, \text{ with } \mathbf{B} \triangleq \mathbf{A}^{-1}. \end{aligned}$$

From Eq. (5.2-9 ??),

$$f_{\mathbf{Y}}(y_1, ..., y_n) = \frac{f_{\mathbf{X}}(x_1^o, ..., x_n^o)}{|J|},$$

where

$$\begin{aligned} |J| &= \text{mag}\left(\det \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial xn} \end{bmatrix}\right) \\ &= \text{mag}\left(\det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}\right) \\ &= |\det \mathbf{A}|. \end{aligned}$$

But, if $\mathbf{B} = \mathbf{A}^{-1}$, then $\det B = 1/\det \mathbf{A}$, and so

$$f_{\mathbf{Y}}(y_1, ..., y_n) = |\det \mathbf{A}| f_{\mathbf{X}}(x_1^o, ..., x_n^o).$$

38. This problem is a special case of problem 5.37 with an easily computable inverse. We cannot solve the original system using the direct method unless we use auxiliary variables

$$\begin{array}{lcl} Y_1 = X_1 + X_2 + \cdots + X_n & & \\ Y_2 = \quad\;\; X_2 + \cdots + X_n & & \Leftarrow \text{ original system} \\ Y_3 = X_3 + \cdots + X_n & & \\ \vdots \quad \vdots \quad\quad \ddots \quad \ddots \quad \vdots \quad \vdots & & \Leftarrow \text{ auxiliary variables} \\ Y_{n-1} = X_{n-1} + X_n & & \\ Y_n = X_n & & \end{array}$$

Now, we have n equations in n unknowns: $\mathbf{y} = (y_1, ..., y_n)^T = \mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & 1 \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix}.$$

$\mathbf{A}^{-1}$ is easily calculated from the n individual equations

$$\begin{aligned}
x_n^o &= y_n \\
x_{n-1}^o &= y_{n-1} - y_n \\
&\vdots \\
x_k^o &= y_k - y_{k+1} \\
&\vdots \\
x_2^o &= y_2 - y_3 \\
x_1^o &= y_1 - y_2.
\end{aligned}$$

Thus

$$\mathbf{A}^{-1} = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 & -1 \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Clearly $\det \mathbf{A} = \det\left(\mathbf{A}^{-1}\right) = 1$, and so

$$\begin{aligned}
f_{\mathbf{Y}}(y_1, ..., y_n) &= |\det \mathbf{A}| f_{\mathbf{X}}(x_1^o, ..., x_n^o) \\
&= f_{\mathbf{X}}(y_1 - y_2, y_2 - y_3, y_3 - y_4, ..., y_n).
\end{aligned}$$

To obtain $f_{\mathbf{Y}}(y_1, y_2)$, we would have to integrate out with respect to $y_3, y_4, ..., y_n$, i.e.

$$f_{\mathbf{Y}}(y_1, y_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{\mathbf{X}}(y_1 - y_2, y_2 - y_3, y_3 - y_4, ..., y_n) dy_3 dy_4 \cdots dy_n$$

Solutions to HW Problems Chapter 6

6.1 The rearrangement leads to $Y_2 = X_1, Y_1 = X_2, Y_4 = X_3, etc.$ so the sequence $X_1, X_2, X_3, ..., X_n$ becomes the sequence $Y_2, Y_1, Y_4, Y_3, ..., Y_n, Y_{n-1}$. Since there is no size ordering or any other functional relationship among the Y's , the elements of the $\{Y_i, i = 1, ..., n\}$ are i.i.d. as well.

6.2 $f_{X_1X_2X_3}(x_1, x_2, x_3) = \prod_{i=1}^{3} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x_i^2\right) = (2\pi)^{-3/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{3} x_i^2\right)$, To get the pdf of $Y \triangleq X_1 + X_2 + X_3$, we need only to remember that the sum of n i.i.d $N(\mu, \sigma^2)$ RVs is distributed as $N(n\mu, n\sigma^2)$. Hence $f_Y(y) = (6\pi)^{-1} \exp\left(-\frac{1}{2}(y/\sqrt{3})^2\right)$.

6.3 Let $X_i = \begin{cases} 1, \text{ if exposed villager dies; } P[X_i = 1] \triangleq p \\ 0, \text{else; } P[X_i = 0] = 1 - p = q \end{cases}$.

Then $\hat{p} = \sum_{i=1}^{n} X_i / n$. Since X_i is Bernoulli, $n\hat{p} = \sum_{i=1}^{n} X_i$ is binomial

with $E[\hat{p}] = p,\ Var[\hat{p}] = p(1-p)/n$. Since n>>1, we use the Normal approximation and write

$$P[-a < (\hat{p} - p) < a] = 0.95 = 2erf\left[\frac{a}{\sqrt{p(1-p)/n}}\right]$$

or $F_{SN}(z_{0.975}) = F_{SN}\left(a\sqrt{n}/\sqrt{p(1-p)}\right) = 0.975$. From the Tables of the SN we get

$z_{0.975} = 1.96$ and $a = 1.96\sqrt{p(1-p)/n}$. Finally we solve $(\hat{p} - p)^2 = 3.84\, p(1-p)/n$ and obtain $p_2 = 0.58, p_1 = 0.46$ and the margin of error is $\pm 6\%$.

6.4 Expand $(p - \hat{p})^2 - (9/n)p(1-p) = 0$ and obtain

$p^2[1 + (9/n)] - p[2\hat{p} + (9/n)] + \hat{p}^2 = 0$,which is of the form $ap^2 + bp + c = 0$; then use the quadratic root formula: $p = -(b/2a) \pm \sqrt{(b/2a)^2 - (c/a)}$ to get the desired result.

6.5 This can be done in Excel as:

1. create a column X that varies from0.05 to 0.95 in steps of 0.05;

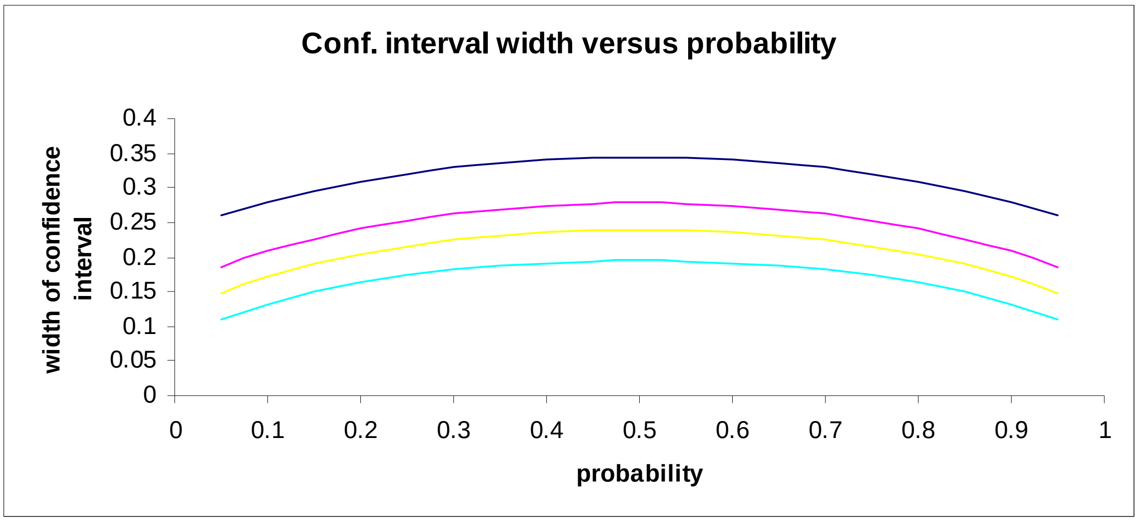
2.Fix the parameter *n*;

3. Create a column as

$$W = 2*SQRT(((2*X+(9/n))/2*(1+(9/n))\wedge 2 - X\wedge 2/(1+(9/n)))$$

4. Create additional columns as you change *n*.

5. Use the Chart Wizard to get the curves below.



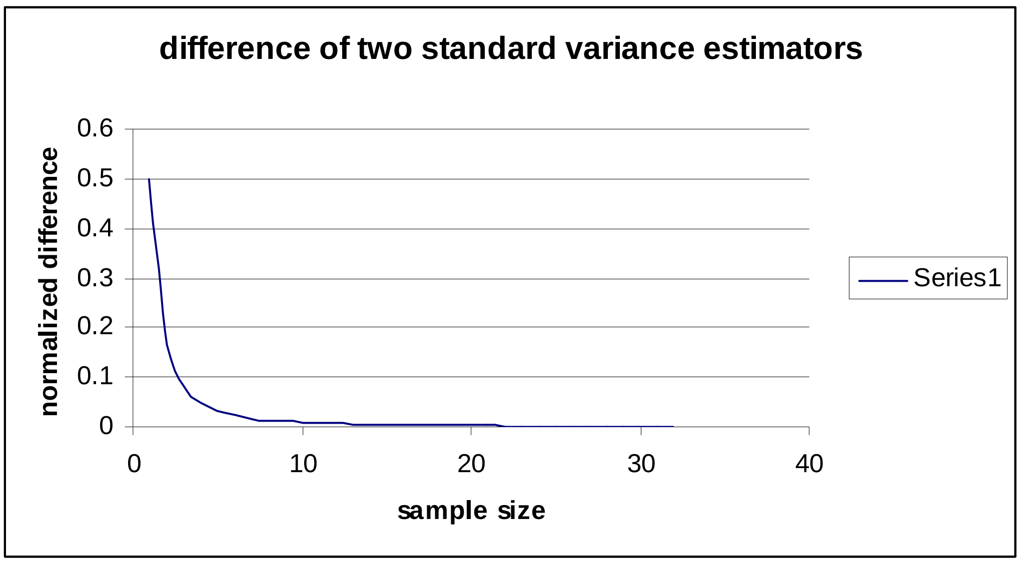
6.6 A reasonable way to do this (here we borrow from the next chapter) is to compute the so-called *distance* from the actual data yield to the *theoretical* data yield if the coin was unbiased. Thus, assume we flip the coin *n* times where $n >> 1$.If the coin is fair then we expect, on the average $n/2$ heads and $n/2$ tails. Let Y denote the observed number of heads; then it is approximately true (by the Central Limit Theorem)

$$\frac{Y-n/2}{\sqrt{n/2}} : N(0,1) \text{ and } \frac{n/2-Y}{\sqrt{n/2}} : N(0,1)$$

and $V \triangleq \left(\frac{Y-n/2}{\sqrt{n/2}}\right)^2 + \left(\frac{n/2-Y}{\sqrt{n/2}}\right)^2$ is a "distance statistic", which is χ^2 with one degree of freedom. Now if a realization of this distance statistic is too large we would reject the notion that that the coin is fair. On the other hand if the distance is small we would accept that the coin is fair. To find the 95 percent cutoff point we go to the tables of the Chi-square distribution and look up the CDF with DOF=1 and $0.95 = F_{\chi^2}(x_{0.95};1)$,which yields $x_{0.95} = 3.84$. Thus if $V > 3.84$ we would reject that the coin is fair.

6.7 Let $\hat{\sigma}_1^2 \triangleq (n-1)^{-1}\sum_{i=1}^{n}\left(X_i - \hat{\mu}_X\right)^2$ and $\hat{\sigma}_2^2 \triangleq n^{-1}\sum_{i=1}^{n}\left(X_i - \hat{\mu}_X\right)^2$. We plot

$(\hat{\sigma}_1^2 - \hat{\sigma}_2^2) / \sum_{i=1}^{n}\left(X_i - \hat{\mu}_X\right)^2 \triangleq$ normalized difference=$1/[(n-1)n]$ using, say, Excel to obtain the figure shown below.



We see that for *n*>20, the difference becomes extremely small.

6.8 Since $X : N(1,1)$ i.e. $\sigma_X = 1$, we can rewrite $P[|\hat{\mu}_X(n) - \mu_X| \leq 0.1]$ as

$$P[\frac{-0.1}{1/\sqrt{n}} \leq \frac{\hat{\mu}_X(n) - \mu_X}{1/\sqrt{n}} \leq \frac{0.1}{1/\sqrt{n}}] = P[-0.1\sqrt{n} \leq Y \leq 0.1\sqrt{n}]$$

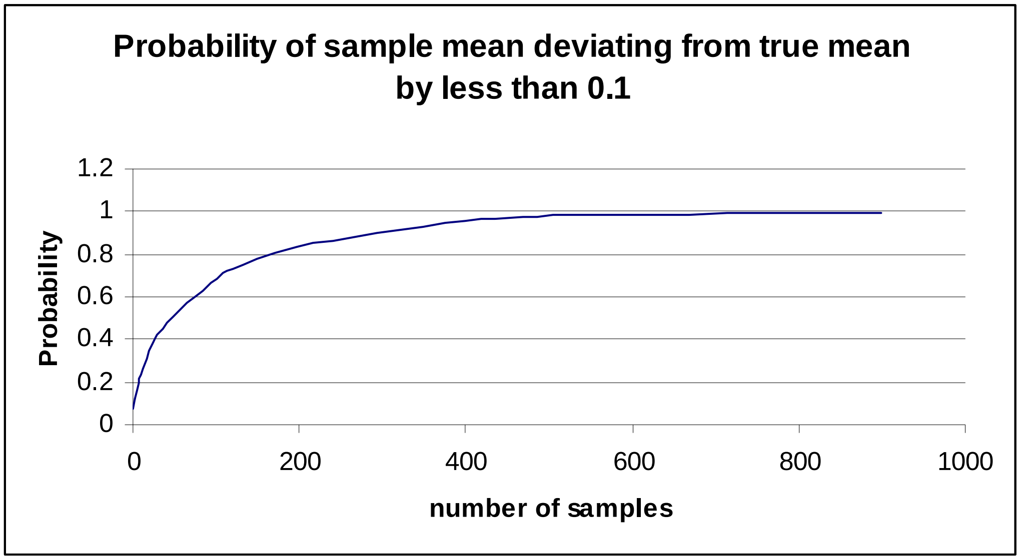
$= 2 \times erf(0.1\sqrt{n})$, where Y is the standard Normal RV.

From Table 2.4-1 we get

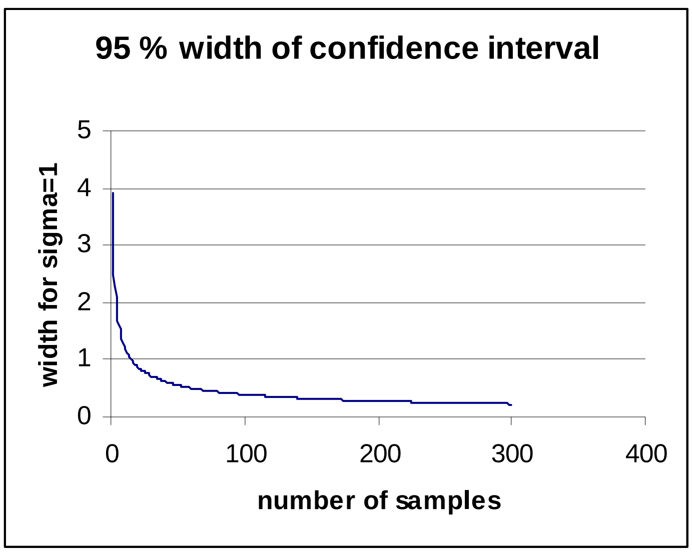
1	0.07966
2.25	0.118
4	0.15852
6.25	0.19742
9	0.23582
25	0.38292
36	0.4515
64	0.57628
100	0.68268

121	0.72866
196	0.83848
400	0.95448
552	0.9812
900	0.99728

Where the first column is n and the second column is P. When plotted we get



6.9 Since the variance is known we can use the formula in Equation 6.3-13: $W_\delta = 2 \times z_{(1+\delta)/2} \times \sigma_X / \sqrt{n}$, which yields, for $\delta = 0.95$ and $\sigma_X = 1 : W_\delta = 3.92 / \sqrt{n}$.



6.10 To simplify things let's temporarily denote $K \triangleq (\alpha!\beta^{\alpha+1})^{-1}$. Then the MGF is given by $M(t) = K\int_0^\infty x^\alpha e^{-x/\beta} e^{tx} dx = K\int_0^\infty x^\alpha e^{-x(1/\beta - t)} dx$. Now let $u \triangleq x(\beta^{-1} - t)$ so that $x = u/(\beta^{-1} - t)$ and $dx = du/(\beta^{-1} - t)$. This transformation yields $M(t) = K \times (1/\beta - t)^{-(\alpha+1)} \int_0^\infty u^\alpha e^{-u} du = K \times (1/\beta - t)^{-(\alpha+1)} \times \alpha!$

Finally, substituting for K enables us to write

$M(t) = (\alpha!)^{-1}\alpha!/[\beta^{(\alpha+1)} \times (\beta^{-1} - t)^{\alpha+1}] = \dfrac{1}{(1-\beta t)^{\alpha+1}}$, which is the desired result.

6.11 There are at least two ways to do this problem: Let $X_1, ..., X_n$ denote the n i.i.d observations on X. (1) Start with $X_1 : N(\mu, \sigma)$ and $Y_1 \triangleq \dfrac{X_1 - \mu}{\sigma}$. Then $Y_1 : N(0,1)$ and let $W_1 \triangleq Y_1^2$. Then $P[W_1 \le w] = F_{W_1}(w) = F_Y(\sqrt{w}) - F_Y(-\sqrt{w})$ or, equivalently, by differentiation, $f_{W_1}(w) = \dfrac{1}{\sqrt{2\pi w}} e^{-w/2} u(w)$. Next, we consider $W_2 \triangleq Y_1^2 + Y_2^2$ and since Y_1^2 and Y_2^2 are i.i.d

$$f_{W_2}(w) = \frac{1}{\sqrt{2\pi w}} e^{-w/2} u(w) * \frac{1}{\sqrt{2\pi w}} e^{-w/2} u(w)$$

$$= \frac{1}{2} e^{-w/2} u(w) \text{ (see Example 4.11).}$$

Proceeding in this way, by repeated convolutions, we arrive at the result:

$$f_{W_n}(w) \triangleq f_W(w;n) = ([n/2) - 1]!2^{n/2})^{-1} w^{(n/2)-1} \exp(-(1/2)w) \times u(w)$$

Another way to obtain the result is to compute the moment generating function of $W_n \triangleq \sum_{i=1}^n Y_i^2$. This is easily done as

$$M_n(t) = (2\pi)^{-n/2} \int_{-\infty}^\infty \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty \exp\left(t\sum_{i=1}^n y_i^2\right) \exp\left(-(1/2)\sum_{i=1}^n y_i^2\right) \prod_{i=1}^n dy_i$$

$$= \left[\int_{-\infty}^\infty \frac{1}{2\pi} \exp\left(-(1/2) y_i^2 (1-2t)\right) dy_i\right]^n = \left(\frac{1}{1-2t}\right)^{n/2}.$$

The MGF of $W_1 \triangleq Y_1^2$ is $M_1(t) = \dfrac{1}{\sqrt{(1-2t)}}\left(\dfrac{1}{(2\pi)^{1/2}}\int_{-\infty}^\infty e^{-\frac{1}{2}y^2} dy\right) = \dfrac{1}{\sqrt{(1-2t)}}$. Hence the MGF of $W_n \triangleq Y_1^2 + Y_2^2 + \cdots Y_n^2$ is $\dfrac{1}{(1-2t)^{1/2}} \times \cdots \times \dfrac{1}{(1-2t)^{1/2}}$ (n times)$= \dfrac{1}{(1-2t)^{n/2}}$.

6.12 The solution to this problem is given in Appendix F. However you might want to try a different approach as follows.

Assume that we have n i.i.d $X_i : N(\mu, \sigma^2), i = 1, ..., n$ RVs. For simplicity let

$Y_i \triangleq \dfrac{X_i - \mu}{\sigma}, i = 1, ..., n$. The joint pdf of the $Y_i,\ i = 1, ..., n$ is $f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-n/2} e^{-\frac{1}{2}\sum_{i=1}^{n} y_i^2}$,where $\mathbf{Y} \triangleq \{Y_i, i = 1, ..., n\}$ and $\mathbf{y} \triangleq \{y_i, i = 1, ..., n\}$.

Now let $m \triangleq \dfrac{1}{n}\sum_{i=1}^{n} y_i$ and $s \triangleq \dfrac{1}{n}\sum_{i=1}^{n}(y_i - \mu_s)^2$. We observe that $\sum_{i=1}^{n} y_i^2 = ns + nm^2$, so that, symbolically at least, we can write

$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-n/2} e^{-\frac{1}{2(1/n)}\left[m^2 + s^2\right]} u(s)$. This form suggests that we define two new RVs $M \triangleq \dfrac{1}{n}\sum_{i=1}^{n} Y_i$, and $S \triangleq \dfrac{1}{n}\sum_{i=1}^{n}(Y_i - M)^2$, which we recognize as the sample mean $\hat{\mu}$ and sample variance $\hat{\sigma}^2$ respectively. Try to compute $f_{MS}(m, s)$ and show that it factors as $f_{MS}(m, s) = f_M(m) f_S(s)$ thus proving the independence of $\hat{\mu}$ and $\hat{\sigma}^2$.

6.13 (a) For simplicity we leave out the subscript n on W_n so that it becomes W. Clearly if X and W are independent so are $Y \triangleq (X - \mu)/\sigma$ and W Now $Y : N(0,1)$ and $W : \chi_n^2$ so that $f_Y(y) = (2\pi)^{-1/2} e^{-\frac{1}{2}y^2}$ and $f_W(w; n) = \dfrac{1}{[(n/2) - 1]! 2^{n/2}} w^{(n/2)-1} e^{-(1/2)w} u(w)$. Now take the product to obtain $f_{YW}(y, w; n) = \dfrac{1}{(2\pi)^{1/2}} e^{-\frac{1}{2}y^2} \dfrac{1}{[(n/2) - 1]! 2^{n/2}} w^{(n/2)-1} e^{-(1/2)w} u(w)$.

(b) We create two new RVs as $T \triangleq \dfrac{Y}{\sqrt{W/n}}$ and $S \triangleq W$ (an auxiliary RV). Now consider the functional transformations:

$t = y/(\sqrt{w/n})$ and $s = w$. The Jacobian is

$$\begin{vmatrix} \dfrac{\sqrt{n}}{\sqrt{s}} & \dfrac{-t}{s} \\ 0 & 1 \end{vmatrix} = \frac{\sqrt{n}}{\sqrt{s}}.$$

Now replacing $t(\sqrt{w/n}) = y$ and $w = s$ in $f_{YW}(y,w;n) = \frac{1}{(2\pi)^{1/2}} e^{-\frac{1}{2}y^2} \frac{1}{[(n/2)-1]!2^{n/2}} w^{(n/2)-1} e^{-(1/2)w} u(w)$ and dividing by $\frac{\sqrt{n}}{\sqrt{s}}$ yields $f(s,t;n) = \frac{1}{(2\pi n)^{1/2}} e^{-\frac{1}{2}s(\frac{t^2}{n}+1)} \frac{1}{[(n/2)-1]!2^{n/2}} s^{(n-1)/2} u(s)$. Finally integrating out with respect to s and recalling that $[(n-1)/2]! \triangleq \int_0^\infty s^{(n-1)/2} e^{-s} ds$ yields the desired result. We use the symbol $[(n-1)/2]! \triangleq \int_0^\infty s^{(n-1)/2} e^{-s} ds$ whether or nor $[(n-1)/2]$ is an integer.

6.14 The RV $F \triangleq \frac{W_m/m}{V_n/n}$ has the F-distribution with m, n degrees of freedom. From the solution of Problem 6.13 we see that $T \triangleq \frac{Y}{\sqrt{W_n/n}}$ has the T-distribution with n degrees of freedom. Now $T^2 \triangleq \frac{Y^2}{W_n/n}$ and $Y : N(0,1)$ so $Y^2 \triangleq W_1$ is χ_1^2. Thus $T^2 \triangleq \frac{W_1/1}{W_n/n} = F_{1,n}$.

6.15 For this problem we used the program Excel.

The first matrix of 20 rows and 6 columns are explained below. Note that in order to save space we show only 20 results instead of the 50 the problem called for.

first column: each cell represent the sample mean of 50 Normal random numbers $\{x_i : 1 = 1, ..., 50\}$ called by the function NORMSINV(RAND());

second column: the sample standard deviation computed from

$$\sigma_s = \left(\frac{1}{49} \sum_{i=1}^{50} \left(x_i - \left[\frac{1}{50} \sum_{i=1}^{50} x_j \right] \right)^2 \right)^{1/2} ;$$

third column: for $\delta = 0.95$, the number $t_{(1+\delta)/2}$: $F_T(t_{(1+\delta)/2}) = (1+\delta)/2$;

fourth column: the lower limit: $\mu_s - (\sigma_s \times t_{(1+\delta)/2}/\sqrt{50})$;

fifth column: the upper limit $\mu_s + (\sigma_s \times t_{(1+\delta)/2}/\sqrt{50})$;

sixth column: interval width $2(\sigma_s \times t_{(1+\delta)/2}/\sqrt{50})$

0.152	1.01	2.01	-0.13514	0.439143	0.574286

0.167	0.984	2.01	-0.11275	0.446751	0.559502
-0.139	0.974	2.01	-0.41591	0.137908	0.553816
0.186	0.871	2.01	-0.06163	0.433625	0.49525
0.023	0.96	2.01	-0.24993	0.295928	0.545856
0.089	0.95	2.01	-0.18108	0.359085	0.54017
-0.027	0.96	2.01	-0.29993	0.245928	0.545856
0.031	1.032	2.01	-0.2624	0.324397	0.586795
-0.133	1.047	2.01	-0.43066	0.164662	0.595324
0.114	1.147	2.01	-0.21209	0.440092	0.652184
-0.038	1.155	2.01	-0.36637	0.290366	0.656733
0.21	1.094	2.01	-0.10102	0.521024	0.622048
-0.07	1.056	2.01	-0.37022	0.230221	0.600441
0.201	0.902	2.01	-0.05544	0.457438	0.512877
0.17	1.065	2.01	-0.13278	0.472779	0.605559
0.193	1	2.01	-0.0913	0.4773	0.5686
0.27	0.874	2.01	0.021522	0.518478	0.496956
0.016	1.137	2.01	-0.30725	0.339249	0.646498
-0.132	1.15	2.01	-0.45894	0.194945	0.65389
0.22	0.97	2.01	-0.05577	0.495771	0.551542
sample mean of 50 samples	**sample sigma**	**t-number for 95%CI**	**lower limit**	**upper limit**	**interval width**

The second matrix shown below repeats the experiment but assumes that the sigma σ_X of the distribution is known to be unity. Again, to save space , we show only 20 outcomes rather than the 50 the problem called for. We n ote that in only one case, marked in bold, did the interval fail to cover the true mean $\mu_X = 0$.

0.028102	1	1.959963	-0.24912	0.305324	0.554445
-0.01337	1	1.959963	-0.29059	0.263852	0.554445
0.246472	1	1.959963	-0.03075	0.523694	0.554445
-0.26262	1	1.959963	-0.53984	0.014602	0.554445
-0.34849	**1**	**1.959963**	**-0.62571**	**-0.07127**	**0.554445**
0.164584	1	1.959963	-0.11264	0.441806	0.554445
0.137643	1	1.959963	-0.13958	0.414865	0.554445
0.130803	1	1.959963	-0.14642	0.408025	0.554445
-0.21185	1	1.959963	-0.48907	0.065372	0.554445
0.037	1	1.959963	-0.24022	0.314222	0.554445
-0.02536	1	1.959963	-0.30258	0.251862	0.554445
-0.04149	1	1.959963	-0.31871	0.235732	0.554445
0.021513	1	1.959963	-0.25571	0.298735	0.554445
0.05607	1	1.959963	-0.22115	0.333292	0.554445
0.214857	1	1.959963	-0.06237	0.492079	0.554445
-0.20623	1	1.959963	-0.48345	0.070992	0.554445
0.064295	1	1.959963	-0.21293	0.341517	0.554445
0.119289	1	1.959963	-0.15793	0.396511	0.554445
0.00526	1	1.959963	-0.27196	0.282482	0.554445

-0.16481	1	1.959963	-0.44203	0.112412	0.554445
sample mean	**given sigma**	**Normal number for 95% CI**	**lower limit**	**upper limit**	**interval width**

Warning: for sensitive and/or important simulations double check that the random number generator indeed gives data that are Normal in distribution.

6.16 We need to show that $E\left[(n-1)^{-1}\sum_{i=1}^{n}\left(X_i-\hat{\mu}_X\right)^2\right]=\sigma_X^2$. This is easily done by expanding the expression in the square bracket as:

$$E\left[(n-1)^{-1}\left(\sum_{i=1}^{n}X_i^2-2n\hat{\mu}_X^2+n\hat{\mu}_X^2\right)\right]$$
$$=E\left[(n-1)^{-1}\left(\sum_{i=1}^{n}X_i^2-n\hat{\mu}_X^2\right)\right]$$
$$=(n-1)^{-1}\left(E[\sum_{i=1}^{n}X_i^2]-nE[\hat{\mu}_X^2]\right)$$
$$=(n-1)^{-1}\left[\left(n\sigma_X^2+n\mu_X^2\right)-nE[\hat{\mu}_X^2]\right]$$

Now

$$E[\hat{\mu}_X^2]=E[\frac{1}{n^2}\left(\sum_{i=1}^{n}X_i^2+\sum_{i=1}^{n}\sum_{j\neq i}^{n}X_iX_j\right)]$$
$$=\frac{1}{n^2}\left(n\mu_X^2+n\sigma_X^2+n(n-1)\mu_X^2\right)$$

so that

$$(n-1)^{-1}\left[\left(n\sigma_X^2+n\mu_X^2\right)-nE[\hat{\mu}_X^2]\right]$$
$$=\frac{1}{n-1}\left(n\sigma_X^2+n\mu_X^2-\sigma_X^2-\mu_X^2-n\mu_X^2+\mu_X^2\right)$$
$$=\sigma_X^2$$

6.17 We need to show that $Var[\hat{\hat{\sigma}}_X^2]\triangleq Var\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\hat{\mu}_X\right)^2\right]\xrightarrow{n\to\infty}0$, where $\hat{\hat{\sigma}}_X^2$ is the biased estimator given by Equation 6.3-4 and the double hat $\hat{}$ is used to differentiate the biased estimator from the unbiased estimator in Equation 6.3-3 i.e., $\hat{\sigma}_X^2\triangleq\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\hat{\mu}_X\right)^2$. However, we have already shown in the text (Section 6.4) that $\hat{\sigma}_X^2$ is consistent for σ_X^2 and $\hat{\sigma}_X^2=k_n\hat{\hat{\sigma}}_X^2$ where $k_n\triangleq\frac{n}{n-1}$. Hence

$k_n\hat{\hat{\sigma}}_X^2$ is consistent for σ_X^2 and since $k_n \triangleq \frac{n}{n-1} \xrightarrow{n\to\infty} 1$, we deduce that $\hat{\hat{\sigma}}_X^2$ is consistent for σ_X^2.

6.18

6.19 From the data we compute that $\hat{\mu}_X = 0.014$ and $\hat{\sigma}_X^2 = 3.89$. We would not be surprised if the data were from $N(0,4)$. In addition, we compute $\sum_{i=1}^{15}\left(X_i - \hat{\mu}_X\right)^2 = 54.45$. Using that the confidence interval is given by

$$\left\{\frac{1}{x_{(1+\delta)/2}}\sum_{i=1}^{15}\left(X_i - \hat{\mu}_X\right)^2, \frac{1}{x_{(1-\delta)/2}}\sum_{i=1}^{15}\left(X_i - \hat{\mu}_X\right)^2\right\}$$

where

$$F_{\chi^2}(x_{(1+\delta)/2};14) = \frac{1+\delta}{2} \text{ and } F_{\chi^2}(x_{(1-\delta)/2};14) = \frac{1-\delta}{2}$$

and $\delta = 0.95$, we compute that that a 95% interval is {2.09, 9.67}. The width of the interval is 7.58. To reduce the width, more data samples would be needed.

6.20 Start with $P[-a \le \frac{\hat{p}-p}{\sqrt{pq/n}} \le a] = \delta$ and let $X \triangleq \frac{\hat{p}-p}{\sqrt{pq/n}}$. Then $P[-a \le X \le a] = 2erf(a) = F_X(a) - F_X(-a) = \delta.$ We recognize that X is $N(0,1)$ under the assumption that the CLT applies. Now a careful examination of the Normal curve and its symmetry reveals that

$F_X(a) + F_X(-a) = 1$ or $F_X(-a) = 1 - F_X(a).$ Thus $2F_X(a) - 1 = \delta$ or $F_X(a) = (1+\delta)/2.$

Solving for a yields $a = x_{(1+\delta)/2}$, the $x_{(1+\delta)/2}$ percentile of X.

6.21 We do this problem using the Excel program. In the first column (A) we list 20 RNs drawn from a $N(0,2)$ population. In B we enter $\max(A_{2n+1}, A_{2n+2})$ for $n = 0,1,...,9.$ In (C) a single estimate of $\hat{\sigma}_X$. The bold number 1.512 in the next-to-last row is the average estimate of $\hat{\sigma}_X$ using Equation 6.4-13. The estimated variance is in the last row at 2.287. The next-to-last entry in column D is the estimated standard deviation using the square root of Equation 6.4-2; it yields a value of 1.505. The associated variance is 2.266. The estimates yielded by the two methods are quite similar.

A	B	C	D
-0.0258	0.5796	2.166806	0.380812
0.5796			1.494506
-0.0132	-0.0132	1.115972	0.396434
-1.592			0.900791
-1.211	-1.211	-1.00695	0.322738
-1.257			0.377119
-0.6477	0.7753	2.513675	2.3E-05
0.7753			2.011291
1.25	1.25	3.355059	3.58307
-2.462			3.309125
-1.284	0.5348	2.0874	0.411009
0.5348			1.386977
-3.521	-2.248	-2.84498	8.28346
-2.248			2.576346
3.237	3.237	6.876923	15.05362
-2.102			2.128973
-0.4927	-0.4927	0.266205	0.02256
-1			0.12752
-1.07	-0.308	0.593577	0.182414
-0.308			0.112158
-0.64289		**1.512369**	**1.505**
		2.287259	**2.266366**

6.22 We write. Then

$$V_1 = (1/3\sigma_X)\times(2X_1 - X_2 - X_3)$$
$$V_2 = (1/3\sigma_X)\times(2X_2 - X_1 - X_3)$$
$$V_1 = (1/3\sigma_X)\times(2X_3 - X_2 - X_1)$$

and

$V_1^2+V_2^2+V_3^2=(1/3\sigma_X^2)\times 2\left(X_1^2+X_2^2+X_3^2-X_1X_2+X_1X_3+X_2X_3\right)$
$=(1/3\sigma_X^2)\left(R_1^2+R_2^2+R_3^2\right)$ where $R_1 \triangleq X_1-X_2, R_2 \triangleq X_1-X_3, R_3 \triangleq X_2-X_3$.
Note that $R_3=R_2-R_1$. Hence
$V_1^2+V_2^2+V_3^2=(2/3\sigma_X^2)\times\left(R_1^2+R_1^2-2R_1R_2\right)$ (so only two RV's are involved!) To get rid of the cross-terms we introduce S_1 and S_2 as
$S_1 \triangleq aR_1+bR_2,\ S_2 \triangleq cR_1+dR_2$ such that
$S_1^2+S_2^2=V_1^2+V_2^2+V_3^2=(2/3\sigma_X^2)\times\left(R_1^2+R_1^2-2R_1R_2\right)$. To compute the coefficients *a, b, c, d* use the procedure in Example 5.5-2.

6.23 The table below summarizes the experiment. In column A we show 20 RNs drawn from $N(0,2)$. Call these $x_i, i=1,...,20$; the bold number in position A21 is the numerical mean $\overline{x}=\frac{1}{20}\sum_{i=1}^{20}x_i$. The numbers in column B are $(x_i-\overline{x})^2, i=1,...,20$ and the bold number in position B21 is $\sum_{i=1}^{20}(x_i-\overline{x})^2$ .Next we find numbers *a,b* such that $F_{\chi^2}(b;19)=0.975$ and $F_{\chi^2}(a;19)=0.025$. This yields *b*=32.9 and *a*=8.91. The last two bold numbers in column A are the lower and upper limits on the 95 percent interval on σ_X^2 . Thus the 95 percent interval is (1.309, 4.833). Another run is shown in columns C and D. Column C is the analogue of A and column D is the analogue of B. There we found the 95 percent confidence interval to be (1.174, 4.338).

A	B	C	D
-0.0258	0.3808	-0.877	0.737228
0.5796	1.494482	0.03777	0.003153
-0.01327	0.396421	0.356	0.14016
-1.592	0.90081	1.198	1.47958
-1.211	0.322749	0.277	0.087249
-1.257	0.377131	2.016	4.138702
-0.6477	2.31E-05	-1.266	1.556556
0.7753	2.011263	1.43	2.097805
1.25	3.583033	-2.99	8.830525
-2.462	3.309161	1.27	1.659923
-1.284	0.411022	-2.17	4.629469
0.5348	1.386954	-0.208	0.035956
-3.521	8.283517	1.66	2.816959
-2.248	2.576378	-0.833	0.663606
3.237	15.05355	-0.4275	0.167379
-2.102	2.129002	-0.2136	0.038111

-0.4927	0.022557	0.5647	0.339982
-1	0.127528	-2.243	4.948934
-1.07	0.182423	2.05	4.278196
-0.308	0.112151	0.001	0.000376
-0.64289	**43.06095**	**-0.01838**	**38.64985**
1.308843	lower limit	**1.174767**	
4.832879	upper limit	**4.337806**	

6.24 We have two i.i.d. RVs X_1 and X_2 and create two new RVs as

$V \triangleq X + (Y/\sqrt{2}\ )$ and $W \triangleq X - (Y/\sqrt{2}\ )$.

Then $E[VW] = E[\left(X + (Y/\sqrt{2}\ )\right)\left(X - (Y/\sqrt{2}\ )\right)] = 1 - 1/2 = 1/2$.

We compute $E[V^2] = E[W^2] = 3/2$. Hence $\rho_{VW} = \dfrac{E[VW]}{\sqrt{\sigma_V^2}\sqrt{\sigma_W^2}} = \dfrac{1/2}{\sqrt{3/2}\sqrt{3//2}} = 1/3$.

To check this out numerically you might call the NORMSINV(RAND()) in Excel to generate two columns of 20 numbers each. Call these columns X_i and Y_i. Next create two new columns of 20 numbers each. Call these columns V_i and W_i and compute $V_i = X_i + Y_i/1.41$ and $W_i = X_i - Y_i/1.41$. Next create column $(V_i - V_s)\times(W_i - W_s)$ where $V_s \triangleq (1/20)\sum_{i=1}^{20} V_i$ and $W_s \triangleq (1/20)\sum_{i=1}^{20} W_i$. Finally compute $\hat{\rho}_{vw} = \sum_{i=1}^{20}(V_i - V_s)\times(W_i - W_s) \Big/ \left(\sqrt{\sum_{i=1}^{20}(V_i - V_s)^2}\sqrt{\sum_{i=1}^{20}(W_i - W_s)^2}\right)$. Do not be alarmed if your result differs from 0.333. The random number generator may not yield properly decorrelated numbers.

6.25 The covariance function is defined as

$c_{11} \triangleq E\left[\left(X - \mu_X\right)(Y - \mu_Y)\right] = E(XY) - \mu_X\mu_Y$.

We wish to show that

$\hat{c}_{11} \triangleq \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu}_X(n))\times(Y_i - \hat{\mu}_Y(n))$ where $\{(X_i Y_i), i = 1,...,n\}$ are paired observations, is an unbiased, consistent estimator for c_{11}.

(a) Unbiasedness: Start with

$$\hat{c}_{11} \triangleq \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i Y_i + \hat{\mu}_X(n)\hat{\mu}_Y(n) - X_i\hat{\mu}_Y(n) - Y_i\hat{\mu}_X(n)\right)$$ and observe that

$$E[\hat{c}_{11}] \triangleq \frac{1}{n-1}\sum_{i=1}^{n}\left(E[X_i Y_i] + E[\hat{\mu}_X(n)\hat{\mu}_Y(n)] - E[X_i\hat{\mu}_Y(n)] - [Y_i\hat{\mu}_X(n)]\right)$$

because of the linearity of the expectation operator.

Now use:

$$E[\hat{c}_{11}] \triangleq \frac{1}{n-1}\sum_{i=1}^{n}\left(E[X_i Y_i] + E[\hat{\mu}_X(n)\hat{\mu}_Y(n)] - E[X_i\hat{\mu}_Y(n)] - [Y_i\hat{\mu}_X(n)]\right)$$

where

$$E[X_i Y_i] = c_{11} + \mu_X \mu_Y$$

$$E[X_i\hat{\mu}_Y(n)] = n^{-1}E[\sum_{i=j}^{n} Y_j X_i] = n^{-1}\left(c_{11} + \mu_X \mu_Y\right) + n^{-1}(n-1)\mu_X \mu_Y$$

$$E[Y_i\hat{\mu}_X(n)] = n^{-1}E[\sum_{i=j}^{n} X_j Y_i] = n^{-1}\left(c_{11} + \mu_X \mu_Y\right) + n^{-1}(n-1)\mu_X \mu_Y$$

$$E[\hat{\mu}_X(n)\hat{\mu}_Y(n)] = n^{-2}E[\sum_{l=1}^{n} X_l \sum_{k=1}^{n} Y_k] = n^{-2}\sum_{l=1}^{n}\sum_{k=1}^{n} E[X_l Y_k]$$

$$= n^{-2}\left(n(c_{11} + \mu_X \mu_Y)\right) + n^{-2}\left(n(n-1)\mu_X \mu_Y\right)$$

Now put all these terms back into the expression for $\hat{c}_{11}$ and find that $E[\hat{c}_{11}] = c_{11}$.

(b) Consistency:

Here we use a "little trick" that mathematicians would not approve of. To prove consistency we have to let *n*, the number of samples get very large. This being the case we replace

$\hat{\mu}_X(n)$ by μ_X and $\hat{\mu}_Y(n)$ by μ_Y since for large *n*, $\hat{\mu}_X(n) \rightarrow \mu_X$ and $\hat{\mu}_Y(n) \rightarrow \mu_Y$.

Also $n-1 \rightarrow n$. Consider then the estimator

$\hat{c}_{11} \triangleq \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_X)\times(Y_i - \mu_Y)$ instead of $\hat{c}_{11} \triangleq \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu}_X(n))\times(Y_i - \hat{\mu}_Y(n))$.

This will save a lot of algebra and yield the same result. Thus

$$E[\hat{c}_{11}^2] = n^{-2}E\left[\sum_{i=1}^{n} \sum_{j=1}^{n}(X_i - \mu_X)(Y_i - \mu_Y)\times(X_j - \mu_X)(Y_j - \mu_Y)\right]$$

$$E[\hat{c}_{11}^2] = n^{-2} E\left[\sum_{j=1}^{n} (X_j - \mu_X)^2 (Y_j - \mu_Y)^2\right]$$
$$+ n^{-2} E\left[\sum_{i=1}^{n} \sum_{j \neq i}^{n} (X_i - \mu_X)(Y_i - \mu_Y) \times (X_j - \mu_X)(Y_j - \mu_Y)\right]$$
$$= n^{-2}\left[n c_{22} + n(n-1) c_{11}^2\right] = \frac{c_{22}}{n} + c_{11}^2 - \frac{c_{11}^2}{n}$$

Finally the variance is

$$\sigma_{\hat{c}_{11}}^2 = E[\hat{c}_{11}^2] - c_{11}^2 = \frac{c_{22} - c_{11}^2}{n} \to 0, \text{ as } n \to \infty.$$

Thus the estimator is consistent as long as c_{22} is finite!

6.26 The geometric distribution is often written as

$P_K(k) = pq^k, k = 0,1,...$ To compute the mean we use

$$\mu = E[K] = \sum_{k=0}^{\infty} k p q^k$$
$$= pq \sum_{k=0}^{\infty} k q^{k-1}$$
$$= pq \frac{d}{dq}\left(\sum_{k=0}^{\infty} q^k\right)$$
$$= pq \frac{d}{dq}\left(\frac{1}{1-q}\right)$$
$$= \frac{q}{1-q}$$

where we used that 1=p+q and $\sum_{k=0}^{\infty} q^k = (1-q)^{-1}$ for $0 < q < 1$. With the help of a little algebra, we can easily derive that $q = \dfrac{\mu}{1+\mu}$ and $p = \dfrac{1}{1+\mu}$ so that $P_K(k) = \dfrac{1}{1+\mu}\left(\dfrac{\mu}{1+\mu}\right)^k$.

To get the variance we first compute

$$E[K^2] = \sum_{k=0}^{\infty} k^2 p q^k$$
$$= \sum_{k=0}^{\infty} \left(k(k-1) + k\right) p q^k$$
$$= pq^2 \sum_{k=0}^{\infty} \left(k(k-1) q^{k-2}\right) + pq \sum_{k=0}^{\infty} k q^{k-1}$$
$$= pq^2 \frac{d^2}{dq^2}\left(\frac{1}{1-q}\right) + pq \frac{d}{dq}\left(\frac{1}{1-q}\right)$$
$$= 2\mu^2 + \mu$$

Hence $Var(K) = E[K^2] - \mu^2 = \mu + \mu^2$.

If $n >> 1$we can use $\hat{\mu} \triangleq \frac{1}{n}\sum_{i=1}^{\infty} K_i$ to get a 95 percent confidence interval on μ. Here the K_i are n i.i.d. observations on K. Using the Central Limit Theorem we approximate $\hat{\mu} : N(\mu, \frac{\mu + \mu^2}{n})$. For a 95 percent interval we seek numbers $-a, a$ such that $P[-a < \frac{\hat{\mu} - \mu}{\sqrt{(\mu + \mu^2)/n}} < a] = 0.95$ or, equivalently, that $F_{\hat{\mu}}(a) = 0.975$. Thus $a = 1.96$ and we solve $\frac{(\hat{\mu} - \mu)^2}{(\mu + \mu^2)/n} = 3.84$. This leads to the quadratic equation $\mu^2(1 - \frac{3.84}{n}) - \mu(2\hat{\mu} + \frac{3.84}{n}) + \hat{\mu}^2 = 0$. Define $\alpha \triangleq 1 - \frac{3.84}{n},\ \beta \triangleq \frac{3.84}{n} + 2\hat{\mu},\ \gamma \triangleq \hat{\mu}^2$; then the roots of the quadratic furnish the upper and lower bounds on the 95 confidence interval.

6.27 The event $\{\zeta : -a \le \frac{\hat{p} - p}{\sqrt{pq/n}} \le a\}$ is clearly identical with the event$\{\zeta : -a\sqrt{pq/n} \le \hat{p} - p \le a\sqrt{pq/n}\}$, which for a>0 is clearly identical with the event
$\{\zeta : (\hat{p} - p)^2 \le a^2 pq/n\}$. Hence $P[-a \le \frac{\hat{p} - p}{\sqrt{pq/n}} \le a] = \delta = P[(\hat{p} - p)^2 \le a^2 pq/n]$.

6.28 The likelihood function for n i.i.d Normal RVs is

$$L(\mu_X) = (2\pi\sigma^2)^{-n/2} \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu_X)^2\right).$$

The log-likelihood function $L'(\mu_X) \triangleq \ln L(\mu)$ yields

$L'(\mu_X) = (-n/2)\ln(2\pi\sigma_X{}^2) - \frac{1}{2\sigma_X{}^2}\sum_{i=1}^{n}(X_i - \mu_X)^2$, which upon differentiating with respect to μ_X, yields the sample mean (mean estimating function):

$\hat{\mu}_X(n) = (1/n)\sum_{i=1}^{n} X_i$. *Squared error consistency* requires that $\lim_{n\to\infty} E[(\hat{\mu}_X(n) - \mu_X)^2] = 0$. Expanding and taking expectations before taking limits yields $E[(\hat{\mu}_X(n) - \mu_X)^2] = \mu_X^2 - (1/n)\mu_X^2 - 2\mu_X^2 + \mu_X^2 + (1/n)\mu_X^2 + (1/n)\sigma_X^2 = (1/n)\sigma_X^2$. Clearly for a finite variance $\lim_{n\to\infty} E[(\hat{\mu}_X(n) - \mu_X)^2] = \lim_{n\to\infty} E[\sigma_X^2 / n] = 0$.

6.29 The likelihood function for *n* i.i.d. exponential RVs is $L(\lambda) = \lambda^n e^{-\lambda\sum_{i=1}^{n} X_i} u(\lambda),\ X_i > 0, i = 1,...,n$. The log-likelihood function yields $L'(\lambda) = n\ln\lambda - \lambda\sum_{i=1}^{n} X_i = n\ln\lambda - \lambda n\hat{\mu}$ and its derivative with respect to λ , set to zero, yields $n/\lambda - n\hat{\mu} = 0$ or $\hat{\lambda}_{ML} = 1/\hat{\mu}$. We note that this value is an extreme point that must maximizes the likelihood function since
$L(0) = L(\infty) = 0$ and $L(\lambda) > 0$ everywhere else.

6.30 Consider *n* i.i.d binomial RVs with $P_{K_i}(k; m, p) \triangleq \binom{m}{k} p^k (1-p)^{m-k}, i = 1,...,n$,

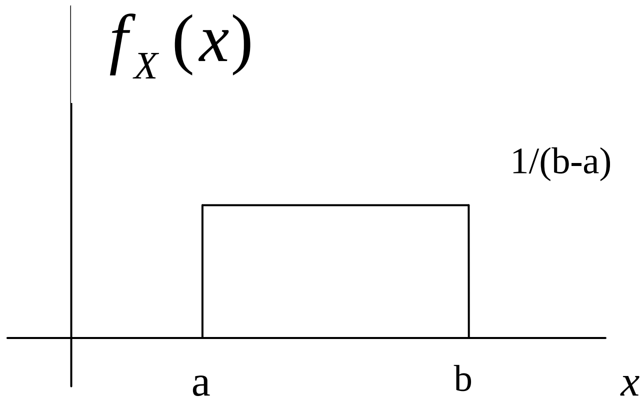
The likelihood function is $L(p) = \binom{m}{K_1}\binom{m}{K_2}\cdots\binom{m}{K_n} p^{\sum_{i=1}^{n} K_i}(1-p)^{\sum_{i=1}^{n}(m-K_i)}$. Then the log-likelihood function is

$L'(p) = \sum_{i=1}^{n}\binom{m}{K_i} + (\sum_{i=1}^{n} K_i)\ln p + (\sum_{i=1}^{n}(m-K_i))\ln(1-p)$. Setting the derivative with respect to *p* to zero yields

$(\sum_{i=1}^{n} K_i)/p\ln p + (\sum_{i=1}^{n}(m-K_i))/(1-p) = 0$ from which we obtain the ML estimator for *p* as

$\hat{p} = \sum_{i=1}^{n} K_i / mn$

6.31 This example illustrates that the ML estimator cannot always be found by differentiation. The pdf of $f_X(x)$ has the shape



The likelihood function for n i.i.d. RVs $X_1, X_2, ..., X_n$ s is $L(a,b) = (b-a)^{-n}$. To maximize $L(a,b)$ we need to make b as small as possible and a as large as possible subject to $b>a$ and the data constraints. Let $X_m \triangleq \min\{X_1, ...X_n\}$. Let $X_M \triangleq \min\{X_1, ...X_n\}$. Clearly $a \le X_m$ and $b \ge X_M$ otherwise realizations would not come from this pdf. The smallest allowable value of b is $b = X_M$; the largest allowable value of a is $a = X_m$. So the MLE of b is $\hat{b}_{ML} = X_M$ and the MLE of a is $\hat{a}_{ML} = X_m$.

6.32 (i) We consider the vector linear model as $\mathbf{Y} = \mathbf{Ia} + b\mathbf{x} + \mathbf{V}$, where $\mathbf{V}$ is a vector of n $N(0,\sigma^2)$ RVs $V_i, i = 1,...,n$. In terms of scalar equations, we write the linear model as:

$Y_i = a + bx_i + V_i,\ i = 1,...,n$. We note that

$E[Y_i] = a + bx_i + E[V_i] = a + bx_i$, since $E[V_i] = 0$ .Also

$\text{Var}[Y_i] = \text{Var}[a + bx_i + V_i] = \text{Var}[V_i] = \sigma^2$. Finally Y_i, exept for a shift, is linearly related to V_i so that $Y_i : N(a + bx_i, \sigma^2)$.

(ii) The likelihood function is

$L(a,b) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left[\frac{Y_i - (a + bx_i)}{\sigma}\right]^2\right)$. So clearly, it is maximized when $l \triangleq \sum_{i=1}^{n}(Y_i - (a + bx)^2$ is minimized. Note that σ^2 is merely a constant here that plays no role in selecting the optimum values of a,b.

(iii) First we compute

$$\frac{dl}{da} = \frac{d}{da}\left(\sum_{i=1}^{n}\left(Y_i - (a+bx_i)\right)^2\right)$$
$$= \sum_{i=1}^{n}\frac{d}{da}\left(Y_i - (a+bx_i)\right)^2$$
$$= \frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{na}{n} - b\frac{1}{n}\sum_{i=1}^{n} x_i = 0$$

With $\hat{\mu}_Y \triangleq \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $\overline{x} \triangleq \frac{1}{n}\sum_{i=1}^{n} x_i$ we get $\hat{a}_{ML} = \hat{\mu}_Y - \hat{b}_{ML}\overline{x}$ .Next we have to compute $\hat{b}_{ML}$ as

$$\frac{dl}{db} = \frac{d}{db}\left(\sum_{i=1}^{n}\left(Y_i - (a_{ML} + bx_i)\right)^2\right)$$
$$= \sum_{i=1}^{n}\frac{d}{db}\left(Y_i - (a_{ML} + bx_i)\right)^2$$
$$= \sum_{i=1}^{n}\left(Y_i x_i - (\hat{\mu}_Y - b\overline{x})x_i - bx_i^2\right)$$
$$= 0$$

Solving for $\hat{b}_{ML}$, we find that

$\hat{b}_{ML} = \dfrac{\sum_{i=1}^{n} Y_i(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$, where we used the identity $\sum_{i=1}^{n}(x_i^2 - \overline{x}^2) = \sum_{i=1}^{n}(x_i - \overline{x})^2$

6.33 We are given that $\mu = 220$ and σ=20 and the weight RV is Normal as $N(220, 400)$. To compute the 95th percentile we solve

$$0.95 = \int_{-\infty}^{x_{0.95}} (800\pi)^{-1/2} \exp[-0.5\left(\frac{x-220}{20}\right)^2]dx$$

which yields $x_{0.95} = z_{0.95}20 + 220 = 1.645 \times 20 + 220 = 253$ lbs . Here $z_{0.95}$ is the 95th percentile of the standard Normal RV.

6.34 We write the geometric distribution as in Table 2.5-1

$F_K(k) = p\dfrac{1-q^{k+1}}{1-q} = 1 - q^{k+1}$ since $p = 1 - q$ .Assume that q>0.5; then we solve 0.5

$1-q^{k_{0.5}+1} \approx 0.5$ or $q^{k_{0.5}+1} \approx 0.5$ or $k_{0.5} \approx \left[\dfrac{\ln 0.5}{\ln q} - 1\right]$,where [ ] is the largest integer not exceeding the contents of the brackets.

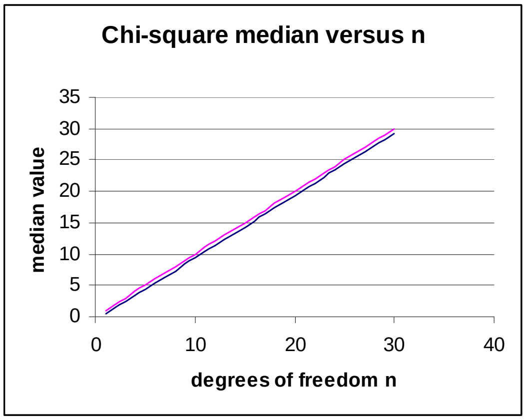
6.35 Let W_n denote the χ_n^2 RV with n degrees of freedom. Then $W_n \triangleq \sum_{i=1}^{n} X_i^2$, where $X_i : N(0,1)$. Clearly $E[W_n] = E\left(\sum_{i=1}^{n} X_i^2\right) = \sum_{i=1}^{n} E[X_i^2] = n$, since $E[X_i^2] = Var[X_i] = 1$. Hence the mean of the χ_n^2 RV with n degrees of freedom is n . To compute the median is a little more difficult. First we compute $Var[W_n]$ as

$$
\begin{aligned}
Var[W_n] &= E[(W_n - E[W_n])^2] = E[W_n^2] - n^2 \\
&= \sum_{i=1}^{n} E[X_i^4] + \sum_{i=1}^{n} [X_i^2] \sum_{j \neq i}^{n} E[X_j^2] - n^2 \\
&= 3n + (n-1)n - n^2 = 2n.
\end{aligned}
$$

There are several ways to compute $\sum_{i=1}^{n} E[X_i^4] = 3n$. The formula

$\int_0^\infty x^{2n} e^{-ax^2} dx = \dfrac{1\cdot 3\cdot 5\cdots(2n-1)}{2^{n+1}a^n}\sqrt{\pi / a}$ is useful with n=2 and a=1/2.

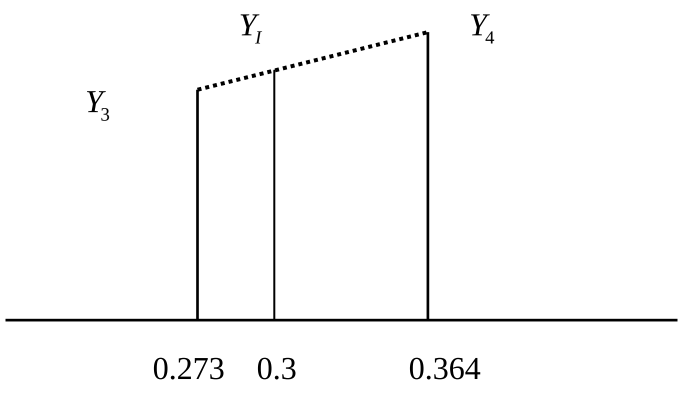
When n is small, the median is easily obtained by table look-up. When n>10 we use the Normal approximation to the χ_n^2 . Then $W_n \sim N(n, 2n)$ and the median is obtained by solving $0.5 = \int_{-\infty}^{x_{0.5}} (4\pi n)^{-1/2} \exp(-0.5\left(\dfrac{x-n}{\sqrt{2n}}\right)^2)dx$. After a conversion to the standard normal we obtain the median as $w_{0.5} \approx n + z_{0.5}\sqrt{2n}$ where $z_{0.5} = 0$, i.e., the 50 percentile point for the standard Normal. A more general formula for the 100u percentile point is $w_u \approx n + z_u\sqrt{2n}$. The approximation (upper curve) is compared with the true median (lower curve) in the figure below.



6.36 Refer to Example 6.8-4. There we showed that

Y_1 estimates $x_{0.0909}$ i.e. the 9th percentile
Y_2 estimates $x_{0.182}$ i.e. the 18th percentile
Y_3 estimates $x_{0.273}$ i.e. the 27th percentile .
Y_4 estimates $x_{0.364}$ i.e. the 36th percentile

Recall that $x_p : F_{SN}(x_p) = p$ is the $100 \times p$ percentile .Now we wish to use the ordered observations $\{Y_i, i = 1, ..., n\}$ to estimate $x_{0.3}$. See the diagram where Y_I is the interpolated value.



Clearly

$$\frac{Y_4 - Y_3}{0.364 - 0.273} = \frac{Y_4 - Y_I}{0.364 - 0.3} \text{ or}$$
$$Y_I = Y_4 + \frac{(0.364 - 0.3) \times (Y_3 - Y_4)}{0.091}$$

6.37 We use the formula $P[Y_1 < x_{0.5} < Y_n] = (1/2)^n \sum_{i=1}^{n-1} \binom{n}{i}$ and find the smallest value of n such that $P[Y_1 < x_{0.5} < Y_n] \geq 0.99$. The table below shows $P[Y_1 < x_{0.5} < Y_n]$ for various values of n:

n	2	3	4	5	6	7	8	9
$P[Y_1 < x_{0.5} < Y_n]$	0.5	0.75	0.88	0.94	0.97	0.98	0.99	0.996

Hence a sample of size will cover the median with probability of 0.99.

6.38 As shown in the text, if $X_1, X_2, \cdots, X_{n-1}, X_n$ are n i.i.d. observations on X with pdf $f(x)$ then their ordered samples $Y_1 < Y_2 < \cdots < Y_n$ have joint pdf

$$f_{\mathbf{Y}}(\mathbf{y}) = n!\, f(y_1) \cdots f(y_n), \text{ where } y_1 < y_2 < \cdots < y_n.$$

Hence $f_{Y_1Y_n}(y_1, y_n) = n!\, f(y_1) f(y_n) \int_{y_1}^{y_{n-1}} \cdots \int_{y_1}^{y_3} f(y_2) \cdots f(y_{n-1}) dy_2 \cdots dy_{n-1}$ i.e. we integrate out with respect to all the y_i except for y_1 and y_n. Now consider the innermost integral (leaving out for simplicity the factor $n!\, f(y_1) f(y_n)$) that is the integration with respect y_2: $\int_{y_1}^{y_3} f(y_2) dy_2 = \int_{y_1}^{y_3} dF = F(y_3) - F(y_1)$. Next consider the integration with respect to y_3:

$\int_{y_1}^{y_4} f(y_3)\big(F(y_3) - F(y_1)\big) dy_3 = \int_{y_1}^{y_4} \big(F(y_3) - F(y_1)\big) dF(y_3)$. If we let $\alpha_i \triangleq F(y_i) - F(y_1)$ then $d\alpha_3 = dF(y_3)$ and

$$\int_{y_1}^{y_4}\left(F(y_3)-F(y_1)\right)dF(y_3)=\int_{\alpha_1}^{\alpha_4}\alpha_3 d\alpha_3=\frac{\alpha_3^2}{2}\Bigg|_{\alpha_1}^{\alpha_4}=\frac{\alpha_4^2}{2}-\frac{\alpha_1^2}{2}=\frac{\alpha_4^2}{2}=\frac{\left(F(y_4)-F(y_1)\right)^2}{2}.$$

Consider next the integration with respect y_4: this yields

$$\int_{y_1}^{y_5}\frac{\left(F(y_4)-F(y_1)\right)^2}{2}f(y_4)dy_4=\int_{y_1}^{y_5}\frac{\left(F(y_4)-F(y_1)\right)^2}{2}dF(y_4)$$

$$\int_{\alpha_1}^{\alpha_5}\frac{\alpha_4^2}{2}d\alpha_4=\frac{\alpha_4^3}{3\cdot 2}\Bigg|_{\alpha_1}^{\alpha_5}=\frac{\alpha_5^3}{3\cdot 2}=\frac{\left(F(y_5)-F(y_1)\right)^3}{3\cdot 2}.$$

It should be clear that after the k^{th} integration we are left with $\frac{\left(F(y_{k+2})-F(y_1)\right)^k}{k!}$. After $n-2$ integrations we have $\frac{\left(F(y_n)-F(y_1)\right)^{n-2}}{(n-2)!}$. Thus the final result (let's not forget the factor $n!\,f(y_1)f(y_n)$) is

$f_{Y_1Y_n}(y_1,y_n)=n(n-1)\left(F(y_n)-F(y_1)\right)^{n-2}f(y_1)f(y_n)$.

6.39 From the previous problem we determined that the joint pdf of $Y_1\triangleq\min(X_1,X_2,...X_n)$ and $Y_n\triangleq\max(X_1,X_2,...X_n)$ is given by

$f_{Y_1Y_n}(y_1,y_n)=n(n-1)\left(F(y_n)-F(y_1)\right)^{n-2}f(y_n)f(y_1)$,

where $f(y)$ and $F(y)$ are the pdf and CDF of X respectively and the $X_i, i=1,...,n$ are i.i.d observations on X. To determine the pdf of the *range* $R\triangleq Y_n-Y_1$ we introduce an auxiliary variable $W\triangleq Y_1$. Then the only solution to the functional equations

$r\triangleq y_n-y_1,\ w\triangleq y_1$ is $y_n=r+w,\ y_1=w$.

The Jacobian magnitude of the transformation is

$$\text{mag}\begin{vmatrix}\partial r/\partial y_1 & \partial r/\partial y_n\\ \partial w/\partial y_1 & \partial w/\partial y_n\end{vmatrix}=1$$

hence

$f_{RW}(r,w)=n(n-1)\left(F(r+w)-F(w)\right)^{n-2}f(r+w)f(w)$, where

$0<r<\infty$ and $-\infty<w<\infty$. To get $f_R(r)$ from $f_{RW}(r,w)$ we integrate out with respect to *w*. This yields

$f(r+w)f(w)=1$.

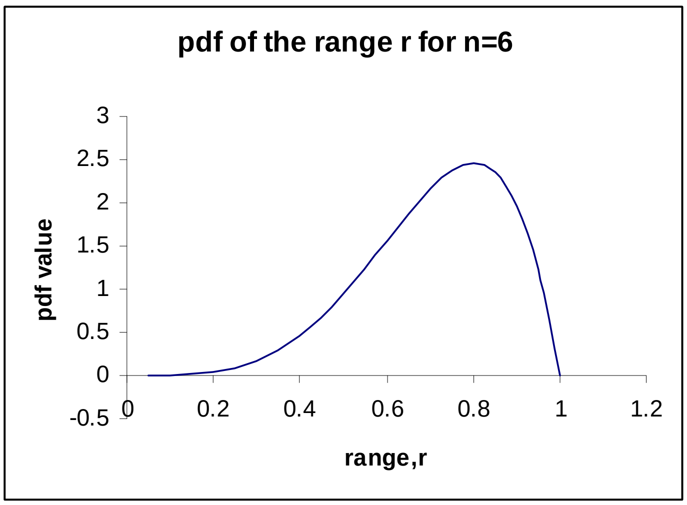
For example, let $f(x)=1,\ 0<x<1,$ and 0 else. Then in the region $0<w<1-r$,

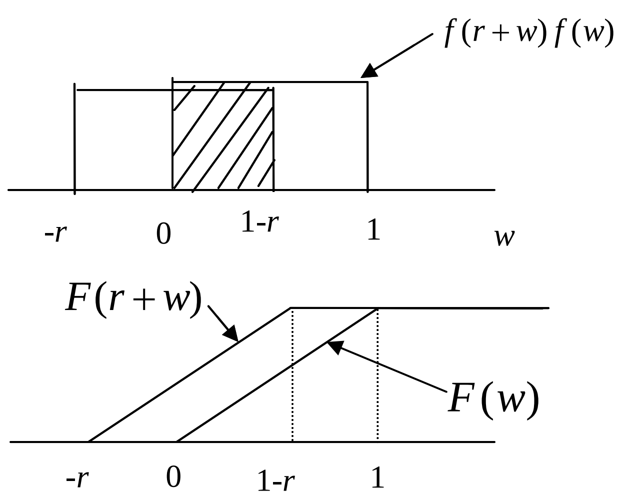
$f(r+w)f(w)=1;$ elsewhere the product is zero. In this region

$F(r+w)-F(w)=(r+w)-w=r$. Hence $f_{RW}(r,w)=n(n-1)r^{n-2}f(r+w)f(w)$ so that

integrating out with respect to w yields

$$f_R(r)=\int_0^{1-r} n(n-1)r^{n-2}dw = n(n-1)(1-r)r^{n-2}\ ,0<r<1.$$





Solutions to HW Problems in Chapter 7

7.1 Begin with

$$B(d) = R(d;\zeta_1)P_1 + R(d;\zeta_2)P_2$$
$$= \left(l(a_1;\zeta_1)P[a_1 \mid \zeta_1] + l(a_2;\zeta_1)P[a_2 \mid \zeta_1]\right)P_1$$
$$+ \left(l(a_1;\zeta_2)P[a_1 \mid \zeta_2] + l(a_2;\zeta_2)P[a_2 \mid \zeta_2]\right)P_2.$$

Next observe that

$$P[a_1 \mid \zeta_2] = \int_c^\infty f(x;\zeta_2)dx;\ P[a_2 \mid \zeta_2] = \int_{-\infty}^c f(x;\zeta_2)dx = 1 - \int_c^\infty f(x;\zeta_2)dx;$$
$$P[a_1 \mid \zeta_1] = \int_c^\infty f(x;\zeta_1)dx;\ P[a_2 \mid \zeta_1] = \int_{-\infty}^c f(x;\zeta_1)dx = 1 - \int_c^\infty f(x;\zeta_1)dx;$$

Inserting these results in the expression for $B(d)$ yields:

$$B(d)\left[l(a_1;\zeta_1)\int_c^\infty f(x;\zeta_1)dx + l(a_2;\zeta_1)\left(1 - \int_c^\infty f(x;\zeta_1)dx\right)\right] \times P_1$$
$$+\left[l(a_1;\zeta_2)\int_c^\infty f(x;\zeta_2)dx + l(a_2;\zeta_2)\left(1 - \int_c^\infty f(x;\zeta_2)dx\right)\right] \times P_2,$$

whence Eq.(7.1-6) follows.

7.2. When $P_1 = 0.9$ and $P_2 = 0.1$ we can expect that cancer is almost always present. We expect the point *c* to move far t the left so that almost any realization of *X* will lead the surgeon to operate. Specializing Eq. (7.1-6) for this case, keeping the loss functions the same as in Table 7.1-1, leads to $B(d) = 31.5 + \int_c^\infty \left(0.5 f(x;\zeta_2) - 30 f(x;\zeta_1)\right)dx$. The Bays solution is when:

$$0.5 f(x;\zeta_2) - 30 f(x;\zeta_1) < 0 \text{ or}$$
$$f(c;\zeta_1) / f(c;\zeta_2) = 0.0167 \quad .$$

Thus when the event $\{X > c\}$ occurs, the surgeon should operate.

7.3 The test is $(n)^{-1}\sum_{i=1}^{n} X_i \triangleq \hat{\mu} > c$ for rejecting the hypothesis. Here c is obtained from

$0.05 = \frac{1}{\sqrt{2\pi\sigma^2/n}}\int_c^{\infty} \exp\left(\frac{-1}{2}\left(\frac{x-\mu_1}{\sigma/\sqrt{n}}\right)^2\right)dx$. When $\sigma = 1$, and converting to the standard, Normal, we get

$0.95 = F_{SN}\left((\sqrt{n}(c-\mu_1)\right) = F_{SN}(1.645)$. Hence $c = \mu_1 + 1.645/\sqrt{n} \to \mu_1$ as $n \to \infty$.

7.4 The power P of a test is $\mathrm{P} = P[\text{reject } H_1 \mid H_2 \text{ true}] = 1 - P[\text{accept } H_1 \mid H_2 \text{ true}]$. Note

$$P[H_2 \text{ true}] = P[\text{reject } H_1, H_2 \text{ true}] + P[\text{accept } H_1, H_2 \text{ true}]$$
$$= P[\text{reject } H_1 \mid H_2 \text{ true}]P[H_2 \text{ true}] + P[\text{accept } H_1 \mid H_2 \text{ true}]P[H_2 \text{ true}]$$

or, equivalently,

$$1 = P[\text{reject } H_1 \mid H_2 \text{ true}] + P[\text{accept } H_1 \mid H_2 \text{ true}]$$

from which it follows that

$$\mathrm{P} = 1 - P[\text{accept } H_1 \mid H_2 \text{ true}].$$

7.5 In Example 7.2-2 it was assumed that X and therefore each of its observation $X_i, i = 1,...,n$ were $N(\mu,\sigma^2)$. The MGF of a Normal RV is $M_X = \exp(\mu t)\times\exp(\sigma^2 t^2/2)$. From the properties of moment generating functions we have that the moment generating function of a sum of i.i.d. RVs is simply the product of their MGFs. Hence

$Y \triangleq \sum_{i=1}^{n} X_i \leftrightarrow \exp(n\mu t)\times\exp(n\sigma^2 t^2/2)$ and Y is seen to be $N(n\mu, n\sigma^2)$. Finally, from elementary probability, the transformation $\hat{\mu} \triangleq Y/n$ yields a $N(\mu,\sigma^2/n)$ RV.

7.6 Here we use the Normal approximation to the binomial since n is large. Under H_1 we find that $\mu = 100\times 0.5 = 50$ and $\sigma^2 = 100\times 0.5\times 0.5 = 25$. Hence we seek the number c such that

$$\alpha = \frac{1}{5\sqrt{2\pi}} \int_k^{\infty} exp\left(-0.5\left((x-50)/5\right)^2\right) dx + \frac{1}{5\sqrt{2\pi}} \int_{-\infty}^{-k} exp\left(-0.5\left((x-50)/5\right)^2\right) dx$$ or, equivalently

$$1-\alpha = 2 \times erf\left(\frac{k-50}{5}\right)$$

At the 0.05 level we find $k = 50 \pm 10$. Hence if either the number of heads or tails exceeds 60 or is less than 40 , the hypothesis is rejected. At the $\alpha = 0.1$ level, we find that $k = 50 \pm (5 \times 1.65)$. Hence if either the number of heads or tails exceeds 58 or is less than 42, the hypothesis is rejected.

7.7 The four decision functions are:

1. Buy the battery if it starts the car, else reject the battery;
2. Buy the battery no matter what;
3. Don't buy the battery no matter what;
4. Don't buy the battery if it starts the car, else buy it.

Put into symbols we get

$$d_1(X): d_1(1) = a_1;\ d_1(0) = a_2$$
$$d_2(X): d_1(1) = a_1;\ d_1(0) = a_1$$
$$d_3(X): d_1(1) = a_2;\ d_1(0) = a_2$$
$$d_4(X): d_1(1) = a_2;\ d_1(0) = a_1$$

We denote the by $\begin{cases} \zeta_1 \text{ the outcome that the battery is of the superior type i.e. from A} \\ \zeta_2 \text{ the outcome that the battery is of the inferior type i.e. from B} \end{cases}$

The state of nature ζ_1 corresponds to battery A with start probability $p_1 = 0.8$. For convenience we write ζ_1=0.8. Likewise the state of nature ζ_2 corresponds to battery B with start probability $p_2 = 0.5$. For convenience we write $\zeta_2 = 0.5$.

The loss functions are: $l(a_1,\zeta_1) = 0;\ l(a_1,\zeta_2) = 40;\ l(a_2,\zeta_1) = 10;\ l(a_2,\zeta_2) = 0.$

The risk formula is: $R(d,\zeta) = l(a_1,\zeta)P[a_1 \mid \zeta] + l(a_2,\zeta)P[a_2 \mid \zeta]$.

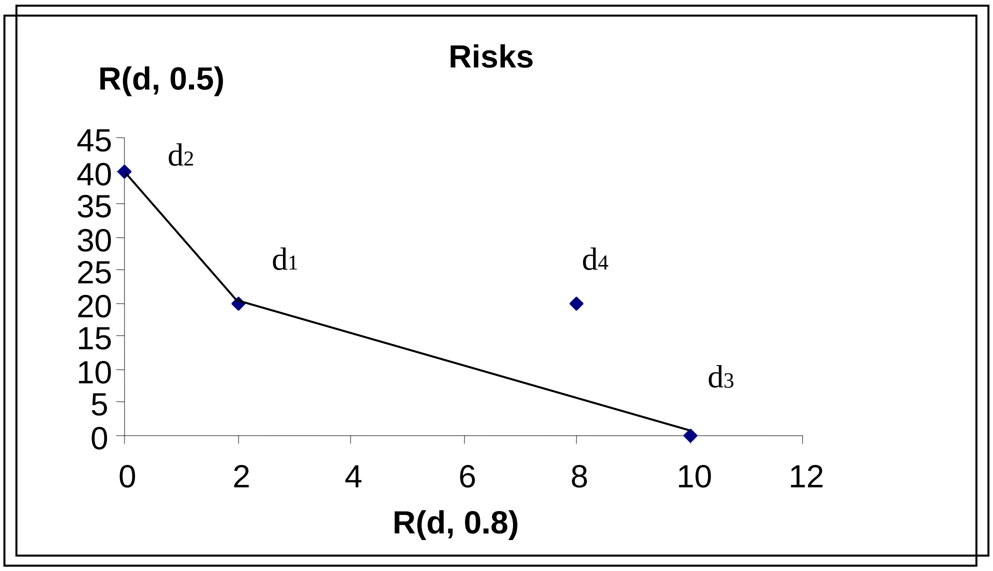
Thus:

$$R(d_1,\zeta_1) = l(a_1,\zeta_1)P[a_1 \mid \zeta_1] + l(a_2,\zeta_1)P[a_2 \mid \zeta_1] = 0 + 10 \times 0.2 = 2$$
$$R(d_1,\zeta_2) = l(a_1,\zeta_2)P[a_1 \mid \zeta_2] + l(a_2,\zeta_2)P[a_2 \mid \zeta_2] = 40 \times 0.5 = 20$$

$$R(d_2,\zeta_1) = l(a_1,\zeta_1)P[a_1 \mid \zeta_1] + l(a_2,\zeta_1)P[a_2 \mid \zeta_1] = 0$$
$$R(d_2,\zeta_2) = l(a_1,\zeta_2)P[a_1 \mid \zeta_2] + l(a_2,\zeta_2)P[a_2 \mid \zeta_2] = 40$$

$$R(d_3,\zeta_1) = l(a_1,\zeta_1)P[a_1 \mid \zeta_1] + l(a_2,\zeta_1)P[a_2 \mid \zeta_1] = 10$$
$$R(d_3,\zeta_2) = l(a_1,\zeta_2)P[a_1 \mid \zeta_2] + l(a_2,\zeta_2)P[a_2 \mid \zeta_2] = 0$$

$$R(d_4,\zeta_1) = l(a_1,\zeta_1)P[a_1 \mid \zeta_1] + l(a_2,\zeta_1)P[a_2 \mid \zeta_1] = 0 + 10 \times 0.8 = 8$$
$$R(d_4,\zeta_2) = l(a_1,\zeta_2)P[a_1 \mid \zeta_2] + l(a_2,\zeta_2)P[a_2 \mid \zeta_2] = 40 \times 0.5 = 20$$



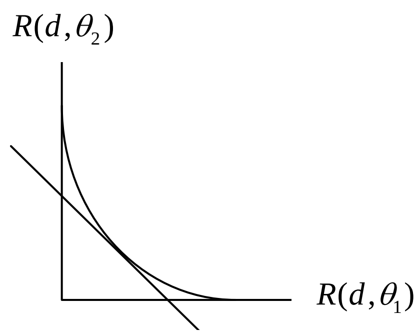
Clearly the decision strategy d_4 is inadmissible.

If $P[A] = 1/3$, and $P[B] = 2/3$ we compute the average risks associated with each strategy as

$$B(d_1) = 1/3 \times 2 + 2/3 \times 20 = 14.3$$
$$B(d_2) = 1/3 \times 0 + 2/3 \times 40 = 26.7$$
$$B(d_3) = 1/3 \times 10 + 2/3 \times 0 = 3.3$$
$$B(d_4) = 1/3 \times 8 + 2/3 \times 20 = 16$$

So clearly, the Bayes strategy is d_3 i.e., don't buy a battery at this shop.

7.8. From the given information, the admissible strategies are on the curve shown below:



$$0.5R(d,\theta_1)+0.5R(d,\theta_2)=c*$$

The value of $c*$ is the point at which they touch. For the curve we have

$\frac{dR(d,\theta_2)}{dR(d,\theta_1)}=-\frac{[R(d,\theta_1)-1]}{[R(d,\theta_2)-1]}$ while for the line we have $\frac{dR(d,\theta_2)}{dR(d,\theta_1)}=-1.$

Hence at the point that they touch $-\frac{[R(d,\theta_1)-1]}{[R(d,\theta_2)-1]}=-1$ and $R(d,\theta_2)=R(d,\theta_1)$. Thus the Bayes strategy corresponds to the decision function for which $R(d,\theta_2)=R(d,\theta_1)$.

.

7.9 We are given $S_1 \triangleq (-\infty,0),\ S_2 \triangleq (0,\infty); \mu_1 = 1/2,\ \mu_2 = -1/2$ and $X : N(\mu,1)$. Hence

$$P[X \in S_1 \,|1/2] = (2\pi)\int_{-\infty}^{0} \exp\left(-0.5(x-1/2)^2\right)dx = F_{SN}(-1/2) = 0.3085$$

$$P[X \in S_2 \,|1/2] = (2\pi)\int_{0}^{\infty} \exp\left(-0.5(x-1/2)^2\right)dx = F_{SN}(1/2) = 0.6915$$

$$P[X \in S_1 \,|-1/2] = (2\pi)\int_{-\infty}^{0} \exp\left(-0.5(x+1/2)^2\right)dx = F_{SN}(1/2) = 0.6915$$

$$P[X \in S_2 \,|-1/2] = (2\pi)\int_{0}^{\infty} \exp\left(-0.5(x+1/2)^2\right)dx = F_{SN}(-1/2) = 0.3085$$

The four decision functions are:

$d_1(X): d_1(X \in S_1) = a_2;\ d_1(X \in S_2) = a_1$
$d_2(X): d_2(X \in S_1) = a_1;\ d_2(X \in S_2) = a_1$
$d_3(X): d_3(X \in S_1) = a_2;\ d_3(X \in S_2) = a_2$
$d_4(X): d_4(X \in S_1) = a_1;\ d_4(X \in S_2) = a_2$

The loss functions are $l(a_1, \mu_1) = 0,\ l(a_1, \mu_2) = 2,\ l(a_2, \mu_1) = 5,\ l(a_2, \mu_2) = 0$

The risk functions are:

$$R(d_1, 1/2) = l(a_1, 1/2)P[a_1 | 1/2] + l(a_2, 1/2)P[a_2 | 1/2] = 0 + 5 \times 0.3085 = 1.54$$

$$R(d_1, -1/2) = l(a_1, -1/2)P[X \in S_2 | 1/2] + l(a_2, -1/2)P[X \in S_1 | 1/2] = 2 \times 0.3085 = 0.61$$

$$R(d_2, 1/2) = l(a_1, 1/2)P[a_1 | 1/2] + l(a_2, 1/2)P[a_2 | 1/2] = 0$$

$$R(d_2, -1/2) = l(a_1, -1/2)P[a_1 | 1/2] + l(a_2, -1/2)P[a_2 | 1/2] = 2$$

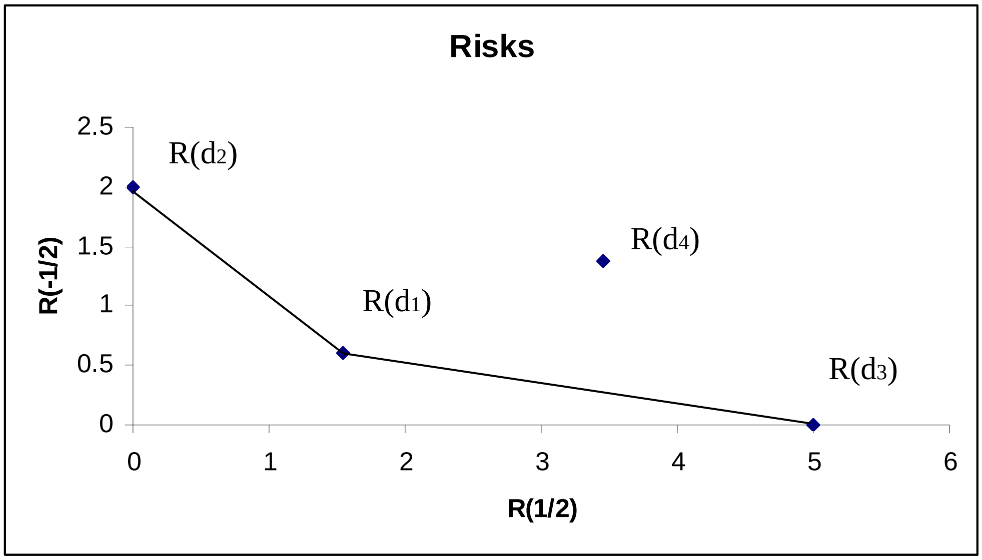
$$R(d_3, 1/2) = l(a_1, 1/2)P[a_1 | 1/2] + l(a_2, 1/2)P[a_2 | 1/2] = 0 + 5 = 5$$

$$R(d_3, -1/2) = l(a_1, -1/2)P[a_1 | 1/2] + l(a_2, -1/2)P[a_2 | 1/2] = 0$$

$$R(d_4, 1/2) = l(a_1, 1/2)P[a_1 | 1/2] + l(a_2, 1/2)P[a_2 | 1/2] = 0 + 5 \times 0.6915 = 3.46$$

$$R(d_4, -1/2) = l(a_1, -1/2)P[a_1 | -1/2] + l(a_2, -1/2)P[a_2 | 1/2] = 2 \times 0.6915 = 1.383.$$

So the four decision functions lead to conditional risks (1.54, 0.61), (0, 2), (5, 0), (3.46, 1.38).



7.10 Assume that we have *m* samples from population X_1 and *n* samples from population X_2. We form the RV

$$T = \frac{\sqrt{mn/(m+n)}\left(\hat{\mu}_1 - \hat{\mu}_2\right)}{\sqrt{\left(\sum_{i=1}^{m}(X_{1i} - \hat{\mu}_1)^2 + \sum_{i=1}^{n}(X_{2i} - \hat{\mu}_2)^2\right)/\left(m+n-2\right)}}$$

and note that H_1, T has the t-distribution with *m*+*n*-2 degrees of freedom. The critical region is the event $T^2 > t_c^2$. To find t_c^2 at the 0.05 level we need to solve P[reject $H_1 \mid H_1$ true]=$P\ [T^2 > t_c^2] = 0.05$, or, equivalently, $P[-t_c < T < t_c] = 0.05$. Hence we find that $F_T(t_{0.025}) = 0.975$ or $t_{0.025} = x_{0.975}$, the 97.5 percentile of the T RV. We find $x_{0.975}$ for a given *m*+*n*-2. For example if *m*=*n*=8, $t_{0.025} = 2.12$. Thus if $T^2 > 4.5$, the hypothesis is rejected.

7.11. The test for $H_1 : \mu_1 = \mu_2$ versus $H_2 : \mu_1 > \mu_2$ the critical region would be $T > t_{0.05} = x_{0.95}$, which we find from $F_T(x_{0.95}) = 0.95$ for the appropriate DOF.

7.12. The test for $H_1 : \mu_1 = \mu_2$ versus H_2: $\mu_1 < \mu_2$ the critical region would be $T < -t_{0.05}$, which we find from $F_T(x_{0.95}) = 0.95$ for the appropriate DOF. Then $-t_{0.05} = -x_{0.95}$

7.13. The test for $H_1 : \sigma^2 = \sigma_0^2$ versus H_2: $\sigma^2 \neq \sigma_0^2$ is done using the Chi-square RV as a statistic i.e. $W_{n-1} \triangleq \sum_{i=1}^{n}(X_i - \hat{\mu}_X)^2/\sigma_0^2$. The critical region is of the form $0 < W_{n-1} < a$ and $b < W_{n-1} < \infty$. An approximate solution for a,*b* is given by $F_{\chi^2}(a) = \alpha/2$ so that $a = x_{\alpha/2}$ i.e. the $x_{\alpha/2}$ percentile of W_{n-1}. Also $F_{\chi^2}(b) = 1 - \alpha/2$ so that

$b = x_{1-\alpha/2}$.

7.14. a) The LRT (for acceptance of the hypothesis) is

$$\Lambda = \frac{(2\pi\sigma^2)^{-1/2}\exp\left(-0.5(X-1)^2/\sigma^2\right)}{(2\pi\sigma^2)^{-1/2}\exp\left(-0.5X^2/\sigma^2\right)} = \exp\left((X-1/2)/\sigma^2\right) > k\,.$$

b) take natural logs of both sides, obtain $X > \sigma^2 \ln k + 1/2 = c$ where $c \triangleq k\sigma^2 + 1/2$.

c) $\alpha = 0.02 = \int_{\infty}^{c} \left(2\pi\sigma^2\right)^{-1/2} \exp\left(-0.5(x-1)^2/\sigma^2\right)dx = F_{SN}(z_{0.02}) = F_{SN}\left((c-1)/\sigma\right)$.

Hence $c = z_{0.02}\sigma + 1$.

d) From the tables we find $z_{0.02} = -2.05$, hence with $\sigma = 1$ we find $c = -1.05$

7.15 The LRT for acceptance of H_1 is

$\Lambda = \dfrac{(2\pi)^{-1/2}\exp\left(-0.5\sum_{i=1}^{n}(X_i-3)^2\right)}{(2\pi)^{-1/2}\exp\left(\left(-0.5\sum_{i=1}^{n}(X_i-1)^2\right)\right)} > k$. Simplifying, we get accept H_1 if $4n\hat{\mu} > \ln k + 8n$, or

if $\hat{\mu} > \dfrac{1}{4n}\ln k + 2 \triangleq c_n$. To find c_n, we solve

$\alpha = P[\text{reject } H_1 \mid H_1 \text{ true}] = F_{SN}\left(\sqrt{n}(c_n-3)\right) = F_{SN}(z_\alpha)$. Thus the main result is

$\sqrt{n}(c_n-3) = z_\alpha$ or $c_n = \dfrac{z_\alpha}{\sqrt{n}} + 3$. For $\alpha = 0.01, z_\alpha = -2.33$ and $c_n = 2.26$.

7.16 We are given $\alpha = 0.02,\ \beta = 0.01$. Hence the power of the test is

$1-\beta = 0.99 = \left(2\pi n^{-1}\right)^{-1/2}\int_{-\infty}^{c_n}\exp\left(-0.5n(x-1)^2\right)dx = F_{SN}\left(\sqrt{n}(c_n-1)\right) = F_{SN}(z_{0.99})$, while from Problem 7.15 we have

$\alpha = P[\text{reject } H_1 \mid H_1 \text{ true}] = F_{SN}\left(\sqrt{n}(c_n-3)\right) = F_{SN}(z_\alpha)$.

Hence we have two equations in two unknowns:

$\sqrt{n}(c_n-1) = z_{0.99} = 2.33$ and $\sqrt{n}(c_n-3) = z_{0.02} = -2.05$.

Solving we get $c_n = 2.2,\ n = 5$ (rounded up from 4.5).

7.17. To keep at $\alpha = 0.02,$ we have to satisfy $c_n = \frac{z_{0.02}}{\sqrt{n}} + 3 = \frac{-2.05}{\sqrt{n}} + 3$. Next, to satisfy a level of power $\text{P}=1-\beta$ we have to satisfy $\sqrt{n}(c_n - 1) = z_\text{P}$. Substituting for c_n and simplifying yields $n = \left(0.5(z_\text{P} + 2.05)\right)^2$. Now go to Excel and create three columns: P, Norm(P,0,1), and *n*. Noted that Norm(P,0,1) returns z_P from which we can compute *n*. Finally call for the Chart Wizard to produce the required graph.

7.18 **m.file:** The following m.file computes realizations of the unbiased sample variances of the two populations P1 and P2.. In this case P1 and P2 are $N(0,1)$

```
function [sigs1,sigs2,lamb]=ftest(n1,n2)
y1=normrnd(0,1,[1,n1]); % this is the Matlab call to generate n1 N(0,1) realizations from P1.
y2=normrnd(0,1,[1,n2]); ]); % this is the Matlab call to generate n2 N(0,1) realizations from P2 .
mu1=sum(y1)/n1; % computes the sample mean of P1
mu2=sum(y2)/n2; % computes the sample mean of P2
z1=(y1-mu1).^2; % this is unnecessary but shows how each point in the P1 data array is squared.
z2=(y2-mu2).^2; % this is unnecessary but shows how each point in the P2 data array is squared.

sigs2=sum((y2-mu2).^2)/(n2-1);
sigs1=sum((y1-mu1).^2)/(n1-1)
```

Command widow:

```
>> [sigs1,sigs2]=ftest(11,21) % this commands calls for 11 samples from P1 and 21 samples from P2
sigs1 =
   0.9855
sigs2 =
   0.9823
>> V=sigs1/sigs2
V =
   1.0033
```

F-test: at the $\alpha = 0.05$ level we find $F_F(x_{0.025};10,20) = 0.025 \rightarrow x_{0.025} = 0.36$ and similarly $F_F(x_{0.975};10,20) = 0.025 \rightarrow x_{0.975} = 2.77$. Since 0.025<V<2.77. we accept the hypothesis that the two populations have the same variance.

7.19 In this problem we have *k* groups (sub-clusters) and the *i*th group $(i = 1,...,k)$ is populated with n_i (assumed continuous) i.i.d. observations $Y_{ij},\ j = 1,...,n_i$. With $Z_i = (1/n_i)\sum_{j=1}^{n_i} Y_{ij} = \hat{\mu}_{Y_i}$ we find that $\mu_{Z_i} = \mu_{Y_i}$ and $\sigma_{Z_i}^2 = \sigma_{Y_i}^2 / n_i$ and if $n_i >> 1$, (assumed true for each i), then $Z_i \xrightarrow{n_i \to \infty} N(\mu_{Z_i}, \sigma_{Z_i}^2)$. The quantity $\sum_{j=1}^{n_i}(Y_{ij} - \hat{\mu}_i)^2$ can be regarded as the *within-group variability* of the ith sub-cluster. The sequence of RVs $Z_i,\ i = 1,...,k$ can then be taken as Normal RVs with (typically) different means and variances. However the sequence of RVs $V_i \triangleq (Z_i - \mu_{Z_i})/\sigma_{Z_i},\ i = 1,...,k$ are $N(0,1)$. Hence $\sum_{i=1}^{k} V_i^2 : \chi_k^2$. Also with $W_i \triangleq (Z_i - \hat{\mu}_{Z_i})/\sigma_{Z_i},\ i = 1,...,k$ we find that $\sum_{i=1}^{k} W_i^2 : \chi_{k-1}^2$, one degree-of-freedom being lost in replacing the mean by the sample mean. The quantity $\hat{\mu}_Z \triangleq (1/k)\sum_{i=1}^{k} Z_i$ is the overall sample mean of the *k* RVs $Z_i, i = 1,...,k$. It represents the "center of gravity" RV of the cluster $\{Z_i,\ i = 1,...,k\}$ and hence of all the data. Note that $\mu_Z \triangleq (1/k)\sum_{i=1}^{k} \mu_{Z_i}$. It is not unreasonable to regard $\sum_{i=1}^{k}(Z_i - \hat{\mu}_Z)^2$ as the RV that measures *inter-group variability* since it measures the total distance-squared from the center-of-gravity of each sub-cluster $\{Y_{ij},\ j = 1,...,n_i\}$ to μ_Z.

7.20 We assume that the within-cluster data $\{Y_{ij},\ j = 1,...,n_i\}$ are i.i.d for each value of *i*. We assume also that between-cluster data are independent. Thus all the data is independent, whether in a specific cluster or not. Now consider the double sum

$S \triangleq \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\dfrac{Y_{ij} - Z_i}{\sigma_{Y_i}}\right)^2$. The inner sum is $\chi_{n_i-1}^2$ so that *S* is the sum of *k* independent *Chi*-square RVs, written, somewhat ingloriously, as $S \triangleq \chi_{n_1-1}^2 + \chi_{n_2-1}^2 + \cdots + \chi_{n_k-1}^2$. The MGF of a Chi-square RV with *m* degrees of freedom is $(1-2t)^{-m/2}$ for $t<1/2$. The MGF of the sum of *k* i.i.d Chi-square RVs is the *k*-product of their MGFs, which in this case is

$(1-2t)^{-(\sum_{i=1}^{k} n_i - k)/2}$ for $t<1/2$. Hence *S* is Chi-square as χ^2_{n-k} where we recalled that $\sum_{i=1}^{k} n_i = n$.

To test for the hypothesis that all the groups are the same i.e. all the data in all the groups come from the same population we construct an F-statistic as follows: *we assume that the* Z_i *are i.i.d.* i.e. $E[Z_i]=\mu_{Y_i} \triangleq \mu_Y$ with $Var[Z_i]=\sigma^2_{Y_i}/n_i=\sigma^2_Y/n_i$. Then $Z_i : N(\mu_Y, \sigma^2_Y/n_i)$ and

$\sum_{i=1}^{k} n_i \left((Z_i-\hat{\mu}_Z)/\sigma_Y\right)^2$ is $\chi^2_{k\text{-}1}$. Then the ratio

$$F_{k-1,n-k} = \left(\sum_{i=1}^{k} n_i \left(Z_i-\hat{\mu}_Z\right)^2/(k-1)\right) \Big/ \left(\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Y_{ij}-Z_i\right)^2/(n-k)\right)$$

is the appropriate statistic is for testing the hypothesis that all the data come from the same population.

7.21 We use the F-test statistic

$$F_{k-1,n-k} = (n-k)\sum_{i=1}^{k} n_i \left(Z_i-\hat{\mu}_Z\right)^2/(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Y_{ij}-Z_i\right)^2,$$

which we rewrite for convenience as

$$F_{k-1,n-k} = \left(\sum_{i=1}^{k} n_i \left(Z_i-\hat{\mu}_Z\right)^2/(k-1)\right) \Big/ \left(\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Y_{ij}-Z_i\right)^2/(n-k)\right)$$

and identify $n=1000$, $n_i=200$, $i=1,...,5$ and $k=5$. From the data we are given

$Z_1=3.17, Z_2=2.72, Z_3=2.63, Z_4=2.29, Z_5=2.19$ yielding $\hat{\mu}_Z=2.6$. Then the numerator is computed as

$200\times\left((3.17-2.6)^2+(2.72-2.6)^2+(2.63-2.6)^2+(2.29-2.6)^2+(2.19-2.6)^2\right)/4=30.225$. To compute denominator, we have to assume that the standard deviations were computed as

$\sigma_i=\left((1/199)\times\sum_{j=1}^{200}(Y_{ij}-Z_i)^2\right)^{1/2}$, $i=1,...,5$. Then it follows that $199\sigma_i^2=\sum_{j=1}^{200}(Y_{ij}-Z_i)^2$, and the denominator is computed as $199\times\left(0.74^2+0.71^2+0.73^2+0.70^2+0.72^2\right)/795=0.65$. Hence the F-statistic is computed as 30.225/0.65=46.5.

Next we go to the F-test calculator e.g. BioKin on line and enter the degrees of freedom $\nu_1 = 4$ (numerator) $\nu_2 = 495$ (denominator) and the significance level 0.05 for a one-sided test and obtain 2.37. Since 46.5>> 2.37 the hypothesis is strongly rejected.

7.22 .We use Pearson's Chi-square test as follows. The expected number of green seeds is $880 \times 0.75 = 660$ while the expected number of yellow seeds is $880 \times 0.25 = 220$. Pearson's statistic yields

$$\Lambda' = \frac{(660-639)^2}{880 \times 0.75} + \frac{(241-220)^2}{880 \times 0.25} = 0.67 + 2.0 = 2.67$$

At the 0.05 level of significance the hypothesis is accepted if $0.001 < \Lambda < 5.02$. Hence the hypothesis is accepted.

7.23. (t-test)

We are given two sets of realizations and told that they come from Normal distributions with the same variance. We use the t-test to test $H_1 : \mu_1 = \mu_2$ versus $H_2 : \mu_1 \neq \mu_2$

Set 1:
-5.980e-1 -9.290e-1 -8.340e-2 1.020e+0 6.780e-1 2.890e-1 1.430e-1 -2.060e+0 1.260e+0 1.670e+0

Set 2:
6.270e-1 2.640e+0 1.530e+0 5.920e-1 1.910e+0 5.050e-1 7.660e-1 2.760e-1 3.070e+0 8.550e-1

Using Excel we compute $\hat{\mu}_1 = 0.14$, $\hat{\mu}_2 = 1.28$. Also writing the t-statistic as *T=NUM/DEN* where $NUM = (\hat{\mu}_1 - \hat{\mu}_2) \times \sqrt{nm/(m+n)}$ and

$DEN = \left(\sum_{i=1}^{m} (X_{1i} - \hat{\mu}_1)^2 + \sum_{j=1}^{n} (X_{2j} - \hat{\mu}_2)^2 \right)^{1/2} / (m+n-2)^{1/2}$ we obtain

$T = NUM / DEN = -2.54 / 1.4 = -1.82$ and $T^2 = 3.31$.This is a two-sided test with 0.025 error probability assigned to each tail. Thus if $T < -t_{0.025}$ or $T > t_{0.025}$ where

$\int_{-\infty}^{t_{\alpha/2}} f_T(x; m+n-2)dx = 1-\alpha/2$. From this we get that $t_{0.025} = 2.1$ At a overall significance of 0.05, the hypothesis is rejected if the t-statistic lies outside the interval (-2.1, 2.1). Since -1.82 is inside this interval, the hypothesis H_1 is accepted.

7.24 Under H_1 we have $p_i = p_{0i}$ all i; hence
$$E[V \mid H_1] = E\left(\sum_{i=1}^{l}(np_{0i})^{-1}(n_i^2 - 2nn_i p_{0i} + n^2 p_i^2)\right)$$
$$= \sum_{i=1}^{l}(np_{0i})^{-1}E(n_i^2 - 2nn_i p_{0i} + n^2 p_i^2)$$
$$= \sum_{i=1}^{l}(np_{0i})^{-1}(np_{0i}(1-p_{0i}) + n^2 p_{0i}^2 - 2n^2 p^2{}_{0i} + n^2 p_{0i}^2)$$
$$= \sum_{i=1}^{l}(np_{0i})^{-1}(np_{0i}(1-p_{0i}) = l-1$$

7.24 **Solution:** Let us compute Λ under the assumption that H_1 is true. Using that $\sum_{i=1}^{m}\left(X_{1i} - \hat{\mu}_1\right)^2 = \sigma^2\chi_{m-1}^2$, $(W_{m-1} : \chi_{m-1}^2)$ and $\sum_{i=1}^{n}\left(X_{2i} - \hat{\mu}_2\right)^2 = \sigma^2\chi_{n-1}^2$ $(W_{n-1} : \chi_{n-1}^2)$ ~~,~~ and factoring out constants we get that

$$\Lambda = A(m,n)\frac{\left(\dfrac{1}{\chi_{m-1}^2 + \chi_{n-1}^2}\right)^{(m+n)/2}}{\left(\dfrac{1}{\chi_{m-1}^2}\right)^{m/2}\left(\dfrac{1}{\chi_{n-1}^2}\right)^{n/2}}, \qquad (7~~6~~.3\text{-}16)$$

where. We recall that the random variable $F_{m,n}$ defined as

$F_{m,n} \triangleq \dfrac{\chi_m^2/m}{\chi_n^2/n}$ is said to have the F-distribution with m and n degrees of freedom respectively (the numerator DOF is cited first). Then rewriting Λ as

$$\Lambda = A(m,n)\frac{\left(\dfrac{n-1}{(n-1)\times\chi_{n-1}^2}\right)^{m/2}\left(\dfrac{1}{\chi_{n-1}^2}\right)^{n/2}\left(\dfrac{1}{[(m-1)/(n-1)\times((n-1)\chi_{m-1}^2/(m-1)\chi_{n-1}^2)]+1}\right)^{(m+n)/2}}{\left(\dfrac{m-1}{(m-1)\times\chi_{m-1}^2}\right)^{m/2}\left(\dfrac{1}{\chi_{n-1}^2}\right)^{n/2}}$$

it follows that

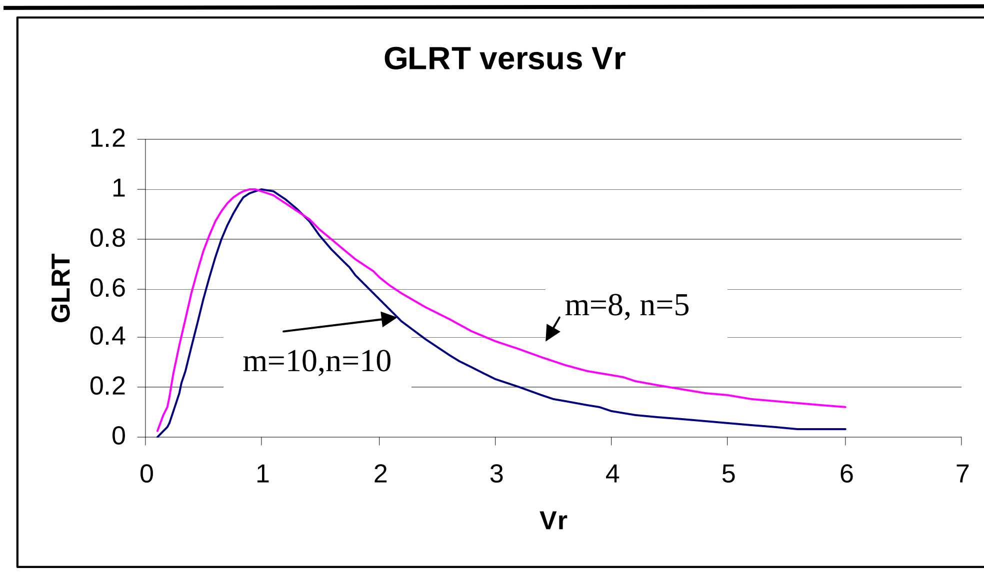
$$\Lambda = A(n,m)\frac{\left(\frac{(n-1)}{(m-1)}F_{n-1,m-1}\right)^{n/2}}{\left(1+\frac{(n-1)}{(m-1)}F_{n-1,m-1}\right)^{(m+n)/2}}$$
$$= A(m,n)\frac{\left(\frac{(m-1)}{(n-1)}F_{m-1,n-1}\right)^{m/2}}{\left(1+\frac{(m-1)}{(n-1)}F_{m-1,n-1}\right)^{(m+n)/2}}.$$

7.25 Under H_2 we have $p_i = p_{1i}$ all i ; hence

$$E[V \mid H_2] = E\left(\sum_{i=1}^{l}(np_{0i})^{-1}(n_i^2 - 2nn_i p_{0i} + n^2 p_{0i}^2)\right)$$
$$= \sum_{i=1}^{l}(np_{0i})^{-1}E(n_i^2 - 2nn_i p_{0i} + n^2 p_{0i}^2)$$
$$= \sum_{i=1}^{l}(np_{0i})^{-1}(np_{1i}(1-p_{1i}) + n^2 p_{1i}^2 - 2n^2 p_{0i} p_{1i} + n^2 p_{0i}^2)$$
$$= \sum_{i=1}^{l}(np_{0i})^{-1}\left((np_{1i}(1-p_{1i}) + n^2(p_{0i} - p_{1i})^2\right).$$

If we differentiate with respect each of the p_{0i} and set the result equal to zero, we find that the minimum occurs when $p_{0i} = p_{1i}$ all i . However, under H_2 this cannot be since at least two of the $p_{0i} \neq p_{1i}$. Clearly $E[V \mid H_2]$ becomes unbounded for p_{0i} close to zero and/or when $n \to \infty$.

7.26 The plot is produced using Excel.



At the 0.05 level this is a two sided test on the RV V_R . We seek the percentile solutions to $F_T(x_{0.025}) = 0.025$ and $F(x_{0.975}) = 0.975$. The solutions are $x_{0.025} = 0.18$ and $x_{0.975} = 9.07$. The test is: reject H_1 if $V_r < x_{0.025} = 0.18$ or $V_R > x_{0.975} = 9.07$

7.27. Most of the proof is given in the text. However, we note that in the text we introduce

$$(m-1)\hat{\sigma}_1^2 = \sum_{i=1}^{m}(X_{1i} - \hat{\mu}_1)^2$$
$$(n-1)\hat{\sigma}_2^2 = \sum_{j=1}^{n}(X_{2j} - \hat{\mu}_2)^2$$

and under H_1, $\sigma_1^2 = \sigma_2^2 = \sigma^2$ so that

$$\begin{aligned}(m-1)\hat{\sigma}_1^2/\sigma^2 &= \sum_{i=1}^{m}\left((X_{1i} - \hat{\mu}_1)/\sigma\right)^2 \triangleq W_{m-1} : \chi^2_{m-1} \\ (n-1)\hat{\sigma}_2^2/\sigma^2 &= \sum_{i=1}^{n}\left((X_{2i} - \hat{\mu}_2)/\sigma\right)^2 \triangleq W_{n-1} : \chi^2_{n-1}\end{aligned}.$$

The variable $F \triangleq \dfrac{(m-1)^{-1}W_{m-1}}{(n-1)W_{n-1}}$ has the F-distribution. Hence $\hat{\sigma}_1^2/\hat{\sigma}_2^2 : F_{m-1,n-1}$.

7.28 The theory behind the test is as follows. The likelihood function is $L(\mu,\sigma^2) = (2\pi\sigma^2)^{n/2}\exp\left(-0.5\left(\sum_{i=1}^{n}[(X_i - \mu)/\sigma]^2\right)\right)$. The global i.e. unrestricted maximum is obtained by differentiating with respect to μ and σ^2. The differentiation yields $\mu^\dagger = \dfrac{1}{n}\sum_{i=1}^{n}X_i = \hat{\mu},\ \sigma^{2\dagger} = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$.

Next, finding the (local) maximum under H_1 yields

$\mu^* = \dfrac{1}{n}\sum_{i=1}^{n}X_i = \hat{\mu},\ \sigma^{2*} = \sigma_0^2$. Taking the ration of $L(\hat{\mu},\sigma_0^2)/L(\hat{\mu},\sigma^{2\dagger}) \triangleq \Lambda$ yields

$\Lambda = (W/n)^{n/2}\exp\left(-0.5(W-n)\right)$, where W under H_1 is Chi-square with DOF n-1. Specifically

$W = \sum_{i=1}^{n}[(X_i - \hat{\mu})/\sigma_0]^2$. A plot of Λ versus W has the appearance as in Figure 7.3-8. Hence acceptance of H_1 requires that $\Lambda > \lambda_C$ or, equivalently, that $w_l < W < w_u$,where w_l, w_u are determined from the type I error criterion and the "equal error rule" discussed in the text. Thus given that $\alpha = P[\text{reject } H_1 \mid H_1 \text{ is true}]$. Thus we seek $\alpha/2 = F_{\chi^2}(w_l; n-1)$ and $1-\alpha/2 = F_{\chi^2}(w_u; n-1)$. Thus we recognize that $w_l = x_{\alpha/2}$ i.e. the $\alpha/2$ percentile point and $w_u = x_{1-\alpha/2}$ i.e. the $1-\alpha/2$ percentile point.

Summary for testing $H_1 : \sigma^2 = \sigma_0^2$ versus $H_2 : \sigma^2 \neq \sigma_0^2$:

1. Obtain realizations $x_1, x_2, ..., x_n$ of $X_1, X_2, ..., X_n$ respectively;

2. Compute the realization of W as $w = \sum_{i=1}^{n}[(x_i - \hat{\mu}')/\sigma_0]^2$, where $\hat{\mu}' = \frac{1}{n}\sum_{i=1}^{n} x_i$ is a realization of $\hat{\mu}$

2. Choose the significance level of the test α e.g. $0.1,\ 0.05,\ 0.025,\ 0.01$;

3. From the tables of the Chi-square CDF find the values $w_l = x_{\alpha/2}$ and $w_u = x_{1-\alpha/2}$ for n-1;

4. If $w_l < w < w_u$, accept H_1, else reject it.

7.29 We use the Pearson test on $H_1:\ P[\text{Heads}]=0.5$ versus $P[\text{Heads}]>0.5$. Hence $V = \left(\frac{35-50\times 1/2}{\sqrt{50\times 1/2}}\right)^2 + \left(\frac{15-50\times 1/2}{\sqrt{50\times 1/2}}\right)^2 = 8$. At the 0.05 level of significance we find $F_{\chi^2}(0.95;1) = 3.84$. Since 8>3.84 we reject $H_1:\ P[\text{Heads}]=0.5$.

7.31 We estimate the 100ith percentile from $\frac{100i}{(n+1)}$. In this case n=24. We note that $y_7 \sim x_{0.28}$, $y_8 \sim x_{0.32}$. Hence $x_{0.3} \sim y_2 + \frac{0.3-0.28}{1/25}(y_3 - y_2) = (y_2 + y_3)/2$. Thus $(y_2 + y_3)/2$ estimates the 30th percentile point.

7.32. We compute for ordered co-joined sequence d=14. From Example 7~~6~~.5-8, the critical value is $d_0 = 6.3$. Since $d_0 < d$ we accept the hypothesis that $\mathrm{P}_1 = \mathrm{P}_2$. Yet it is obvious that P_1 generates even numbers while P_2 generates odd numbers. The run test is not sensitive to populations with all even/odd parity.

7.33.It is clear from the data that if the population generating the S_1 data is X and the population generating the S_2 data is Y then $Y = -10 \times X$. So the correlation coefficient is -1. So in this sense the source are the same since given Y you can get X. However since d=2, the run test result would say that the sources are different.

7.34 We compute for ordered co-joined sequence d=14. From Example 7~~6~~.5-8, the critical value is $d_0 = 6.3$. Since $d_0 < d$ we accept the hypothesis that $\mathrm{P}_1 = \mathrm{P}_2$. Yet it is obvious that P_1 generates even numbers while P_2 generates odd numbers. The run test is not sensitive to populations with all even/odd parity.

Chapter 8 solutions

1. We recall that the set $\{A_n\}_{n=1}^N = \{A_1, A_2, ..., A_N\}$, and wish to prove the chain rule

$$P[A_1A_2\cdots A_N] = P[A_1]P[A_2|A_1]\cdots P[A_N|A_1A_2\cdots A_{N-1}]. \tag{1}$$

We can use mathematical induction in this case. By definition of conditional probability, we know

$$P[A_1A_2] = P[A_1]P[A_2|A_1], \tag{2}$$

thus the proposition is true for $N = 2$. Following mathematical induction, we assume that (1) is true for some positive integer $N = K$, for simplicity denoting the joint event $B \triangleq A_1A_2\cdots A_K$, then by using (2), we have

$$\begin{aligned}
P[A_1A_2\cdots A_{K+1}] &= P[BA_{K+1}] \\
&= P[B]P[A_{K+1}|B] \\
&= P[A_1A_2\cdots A_K]P[A_{K+1}|A_1A_2\cdots A_K] \\
&= P[A_1]P[A_2|A_1]\cdots P[A_K|A_1A_2\cdots A_{K-1}]P[A_{K+1}|A_1A_2\cdots A_K],
\end{aligned}$$

since (1) is assumed true for K. Thus we have shown that (1) being true for $N = K$ implies that it is true also for $N = K + 1$. Thus by the principle of mathematical induction, since we also know (1) is true for $N = 2$, it must be true for all positive integers N.

Expressing this result in terms of joint CDFs $F(x_1, x_2, ..., x_N)$, we have

$$F(x_1, x_2, ..., x_N) = F(x_1)F(x_2|x_1)\cdots F(x_N|x_1, x_2, ..., x_{N-1}).$$

For joint pdf's $f(x_1, x_2, ..., x_N)$, we similarly have

$$f(x_1, x_2, ..., x_N) = f(x_1)f(x_2|x_1)\cdots f(x_N|x_1, x_2, ..., x_{N-1}).$$

2. Given the $N-$dimensional vector $(x_1, x_2, ..., x_N)$ whose components are *pairwise independent*,

$$\text{i.e.} \quad f(x_i, x_j) = f(x_i)f(x_j) \quad \text{for all } i \neq j,$$

we want to show that it is possible that,

$$f(x_1, x_2, ..., x_N) \neq f(x_1)f(x_2)\cdots f(x_N)$$

i.e. *joint independence* does not follow. Consider a case with $N = 3$: $f(x_3, x_2, x_1)$. By the chain rule for pdf's we then have $f(x_3, x_2, x_1) = f(x_3|x_2, x_1)f(x_2|x_1)f(x_1)$ and from pairwise independence we have $f(x_2, x_1) = f(x_2)f(x_1)$, $f(x_3, x_1) = f(x_3)f(x_1)$, and $f(x_3, x_2) = f(x_3)f(x_2)$, substituting in, we conclude

$$f(x_3, x_2, x_1) = f(x_3|x_2, x_1)f(x_2)f(x_1).$$

The question is now whether $f(x_3, x_1) = f(x_3)f(x_1)$, and $f(x_3, x_2) = f(x_3)f(x_2)$ provide enough information to conclude $f(x_3|x_2, x_1) = f(x_3)$. Alas, this is not so. [1]

[1] There is one exception to this and that is the case where the RVs are jointly Gaussian distributed.

Here is a specific counterexample: Let X_1 and X_2 be two independent RVs, each uniformly distributed on the interval $[-\pi, +\pi]$, i.e. $X_i : U[-\pi, +\pi], i = 1, 2$. In terms of pdf's, we have

$$f_{X_i}(x_i) = \begin{cases} \frac{1}{2\pi}, & |x_i| \le \pi, \\ 0, & \text{else.} \end{cases}$$

Next, define a third RV by $X_3 \triangleq (X_1 + X_2) \bmod \pi$, meaning

$$X_3 = \begin{cases} X_1 + X_2 - 2\pi, & X_1 + X_2 > \pi, \\ X_1 + X_2, & |X_1 + X_2| \le \pi, \\ X_1 + X_2 + 2\pi, & X_1 + X_2 < -\pi. \end{cases}$$

Upon some reflection, we see

$$f_{X_3|X_1}(x_3|x_1) = \frac{1}{2\pi}, \quad |x_3| \le \pi,$$

and the same for $f_{X_3|X_2}$, and thus since X_1 and X_2 are independent, we can conclude that X_1, X_2, X_3 are pairwise independent. However, by the definition of X_3, we see that $(X_1 + X_2) \bmod \pi$ determines X_3, specifically

$$f_{X_3|X_1,X_2}(x_3|x_1, x_2) = \delta(x_3 - (x_1 + x_2) \bmod \pi).$$

Thus, joint independence does not prevail.

3. We are given $X_i = X_{i-1} + B_i = \sum_{j=1}^{i} B_j, \ \ 1 \le i \le 5.$

   (a) Thus

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

   (b) Writing $\mathbf{B}^T = (B_1, B_2, B_3, B_4, B_5)^T$, we have $\boldsymbol{\mu}_{\mathbf{X}} = E[\mathbf{AB}] = \mathbf{A}\boldsymbol{\mu}_{\mathbf{B}} = \frac{1}{2}\mathbf{A1}$ where $\mathbf{1}$ is a column vector of all 1s.

   (c) $\mathbf{K_B} = E[\mathbf{B}_c\mathbf{B}_c^T]$ where $\mathbf{B}_c \triangleq \mathbf{B} - \boldsymbol{\mu}_{\mathbf{B}} = \mathbf{B} - \frac{1}{2}\mathbf{1}$. Now

$$(B_c)_i = \begin{cases} +\frac{1}{2}, & p = \frac{1}{2}, \\ -\frac{1}{2}, & p = \frac{1}{2}, \end{cases}$$

thus $E[(B_c)_i^2] = \frac{1}{4}$ and $E[(B_c)_i\,(B_c)_j] = 0$ for $i \ne j$, thus

$$\mathbf{K_B} = \frac{1}{4}\mathbf{I}.$$

(d)

$$
\begin{aligned}
\mathbf{K_X} &= E[(\mathbf{A}\mathbf{B}_c)\,(\mathbf{A}\mathbf{B}_c)^T] \\
&= \mathbf{A}E[\mathbf{B}_c\mathbf{B}_c^T]\mathbf{A}^T \\
&= \mathbf{A}\mathbf{K_B}\mathbf{A}^T \\
&= \frac{1}{4}\mathbf{A}\mathbf{A}^T \\
&= \frac{1}{4}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}.
\end{aligned}
$$

4. (a) Yes. The θ_k are then N outcomes ς (zeta) in the sample space $\Omega = \{\theta_k\}_{k=0}^{N-1}$. The field $\mathcal{F}$ of events is the collection of all 2^N subsets of Ω. The P measure of the event $E \in \mathcal{F}$ is $P[E] = n/N$ where $n = \#$ of outcomes in E. The random variables $X(n, \varsigma)$ map the sample space Ω into the linear space of real-valued sequences.

(b)

$$
\begin{aligned}
E[X[n]] &= E\left[\cos\left(\frac{2\pi n}{5} + \Theta\right)\right] \\
&= \frac{1}{N}\sum_{k=0}^{N-1}\cos\left(\frac{2\pi n}{5} + \frac{2\pi k}{N}\right) \\
&= \cos\left(\frac{2\pi n}{5}\right)\left(\frac{1}{N}\sum_{k=0}^{N-1}\cos\left(\frac{2\pi k}{N}\right)\right) - \sin\left(\frac{2\pi n}{5}\right)\left(\frac{1}{N}\sum_{k=0}^{N-1}\sin\left(\frac{2\pi k}{N}\right)\right) \\
&= \cos\left(\frac{2\pi n}{5}\right)\cdot 0 - \sin\left(\frac{2\pi n}{5}\right)\cdot 0 \\
&= 0.
\end{aligned}
$$

Here, we first used the trigonometry formula for the cosine of the sum of two angles $\cos(A+B)$ and then made indirect use of the formula for the sum of a geometric series, in this case powers of $e^{\frac{j2\pi}{N}}$, as:

$$
\begin{aligned}
\sum_{k=0}^{N-1} e^{\frac{j2\pi k}{N}} &= \sum_{k=0}^{N-1}\cos\left(\frac{2\pi k}{N}\right) - j\left(\frac{1}{N}\sum_{k=0}^{N-1}\sin\left(\frac{2\pi k}{N}\right)\right) \\
&= \frac{1 - e^{j2\pi}}{1 - e^{\frac{j2\pi}{N}}} = 0 + j0.
\end{aligned}
$$

(c)

$$
\begin{aligned}
E[X[n]X[m]] &= E\left[\cos\left(\frac{2\pi n}{5}+\Theta\right)\cos\left(\frac{2\pi m}{5}+\Theta\right)\right]\\
&= \frac{1}{N}\sum_{k=0}^{N-1}\cos\left(\frac{2\pi n}{5}+\frac{2\pi k}{N}\right)\cos\left(\frac{2\pi m}{5}+\frac{2\pi k}{N}\right)\\
&= \frac{1}{2}\cos\left(\frac{2\pi(n-m)}{5}\right)\cdot\left(\frac{1}{N}\sum_{k=0}^{N-1}\cos(0)\right)+\frac{1}{2}\cos\left(\frac{2\pi(n+m)}{5}\right)\cdot\left(\frac{1}{N}\sum_{k=0}^{N-1}\cos\left(\frac{4\pi k}{N}\right)\right)\\
&= \frac{1}{2}\cos\left(\frac{2\pi(n-m)}{5}\right)\cdot 1+\frac{1}{2}\cos\left(\frac{2\pi(n+m)}{5}\right)\cdot 0\\
&= \frac{1}{2}\cos\left(\frac{2\pi(n-m)}{5}\right),
\end{aligned}
$$

where we have made use of the trigonometry formula $\cos A\cdot\cos B$ and also the geometric sum

$$
\begin{aligned}
\sum_{k=0}^{N-1} e^{\frac{j4\pi k}{N}} &= \sum_{k=0}^{N-1}\cos\left(\frac{4\pi k}{N}\right)-j\left(\frac{1}{N}\sum_{k=0}^{N-1}\sin\left(\frac{4\pi k}{N}\right)\right)\\
&= \frac{1-e^{j4\pi}}{1-e^{\frac{j4\pi}{N}}}=0+j0,\quad \text{for } N>2.
\end{aligned}
$$

5. We know that an RV X is defined as a mapping from a sample space Ω to the real line R^1, written symbolically as

$$X:\Omega\longrightarrow R^1,$$

with field of events $\mathcal{F}$ and probability measure P assigned to these events. Let us assume in this case that our sample space is a copy the real line itself R^1. We can then write $X: R^1 \longrightarrow R^1$ or $X(x)=x$ for all $x\in R^1$. Let the field be the Borel field of subsets of R^1 generated by the intervals $(a,b]$ for all $a<b$ which are half-open half-closed. Then any interval in R^1 can be given (represented) by using countable intersections of these intervals or their complements. For example: $(a,\infty)=(-\infty,a]^c$ and $(a,b)=\lim_{n\longrightarrow\infty}(a,b-\frac{1}{n}]$. We can also assign a probability to these events (in this case intervals) and therefore all together we have created an underlying probability space $(\Omega,\mathcal{F},P)$. To do this we use the CDF of the RV X and write

$$P[(a,b]]=F_X(b)-F_X(a).$$

6. (a) Let events S_1 and S_2 be defined as follows for two times $t_2>t_1>0$:

$$
\begin{aligned}
S_1 &\triangleq \{\text{no photon emitted prior to time } t_1\}\\
S_2 &\triangleq \{\text{ at least one photon emitted prior to time } t_2\}.
\end{aligned}
$$

By definition

$$
\begin{aligned}
P[S_2|S_1] &= \frac{P[S_2S_1]}{P[S_1]} \text{ and}\\
P[S_1] &= 1-\int_0^{t_1}\lambda e^{-\lambda t}dt=e^{-\lambda t_1}.
\end{aligned}
$$

Thus

$$P[S_2S_1] = \int_{t_1}^{t_2} \lambda e^{-\lambda t} dt = e^{-\lambda t_1} - e^{-\lambda t_2}$$

and so

$$\begin{aligned} P[S_2|S_1] &= \frac{e^{-\lambda t_1} - e^{-\lambda t_2}}{e^{-\lambda t_1}} \\ &= 1 - e^{-\lambda(t_2 - t_1)}. \end{aligned}$$

(b) Let us define four events as follows:

$$\begin{aligned} A &\triangleq \{\text{at least one photon emitted prior to time } t_2 \text{ from 3 independent sources}\}, \\ S_1 &\triangleq \{\text{no photon emitted from source 1 prior to time } t_2\}, \\ S_2 &\triangleq \{\text{no photon emitted from source 2 prior to time } t_2\}, \quad \text{and} \\ S_3 &\triangleq \{\text{no photon emitted from source 3 prior to time } t_2\}. \end{aligned}$$

Then $P[A] = 1 - P[S_1S_2S_3]$, and because the three sources are independent $P[S_1S_2S_3] = P[S_1]P[S_2]P[S_3]$. Furthermore $P[S_i] = 1 - \int_0^{t_2} \lambda e^{-\lambda t} dt = e^{-\lambda t_2}$. Thus

$$\begin{aligned} P[A] &= 1 - P[S_1]P[S_2]P[S_3] \\ &= 1 - e^{-3\lambda t_2}. \end{aligned}$$

7. (a) We use the general result $E[X] = E[[X|Y]]$. In this instance, it becomes

$$E[e^{j\omega X}] = E[E[e^{j\omega X}|M]].$$

Now $E[e^{j\omega X}|M = m] = \exp(j\omega m - \frac{1}{2}\sigma^2\omega^2)$. Therefore the characteristic function for X can be written as

$$\begin{aligned} \Phi_X(\omega) &= E[e^{j\omega X}] \\ &= E[\exp(j\omega M - \frac{1}{2}\sigma^2\omega^2)] \\ &= e^{-\frac{1}{2}\sigma^2\omega^2} E[\exp j\omega M] \\ &= e^{-\frac{1}{2}\sigma^2\omega^2} \Phi_M(\omega). \end{aligned}$$

Now $\Phi_M(\omega) = E[e^{j\omega M}] = \int_0^\infty e^{j\omega m} \lambda e^{-\lambda m} dm = \lambda/(\lambda - j\omega)$. Thus

$$\Phi_X(\omega) = \frac{\lambda e^{-\frac{1}{2}\sigma^2\omega^2}}{\lambda - j\omega}.$$

(b) For the mean we write

$$\begin{aligned} E[X] &= E[E[X|M]] \\ &= E[M], \end{aligned}$$

i.e. the mean of X is the mean of M, and for the variance

$$\begin{aligned} \sigma_X^2 &= E[X^2] - \mu_X^2 \\ &= E[\Sigma^2 + M^2] - \mu_X^2 \\ &= \mu_{\Sigma^2} + E[M^2] - \mu_X^2 \\ &= \mu_{\Sigma^2} + \sigma_M^2. \end{aligned}$$

8. (a) First we define the dummy RV $S \triangleq X+Y$, then we have $X = SR$ and $Y = S - SR$. Then $f_{R,S}(r,s) = f_{X,Y}(x,y)|J|$ for $0 \le r \le 1, 0 \le s < \infty$, and the Jacobian J is given as

$$J = \det \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial s} \end{bmatrix} = \det \begin{bmatrix} s & r \\ -s & 1-r \end{bmatrix} = s.$$

Since X and Y are independent

$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x) f_Y(y) \\ &= \lambda^2 e^{-\lambda(x+y)} \quad \text{(by the problem statement)} \\ &= \lambda^2 e^{-\lambda s}. \end{aligned}$$

So $f_{R,S}(r,s) = \lambda^2 s e^{-\lambda s}$. Then integrating out the variable s, we get

$$\begin{aligned} f_R(r) &= \int_0^\infty \lambda^2 s e^{-\lambda s} ds \\ &= 1. \end{aligned}$$

Remembering that the range of r is $[0,1]$, we have finally

$$f_R(r) = \begin{cases} 1, & 0 \le r \le 1, \\ 0, & \text{else.} \end{cases}$$

(b) Since $A \triangleq \{X < 1/Y\}$, we can write

$$\begin{aligned} P[X \le x | A, Y = y] &= \frac{P[X \le \min(x, 1/y), Y = y]}{P[A, Y = y]} \quad \text{(since the RVs are continuous)} \\ &= \frac{P[X \le \min(x,1/y)] P[Y = y]}{P[Y = y] P[A | Y = y]} \\ &= \frac{P[X \le \min(x, 1/y)]}{P[A|Y = y]} \quad \text{(cancelling like terms }^2\text{)} \\ &= \frac{P[X \le \min(x,1/y)]}{P[X \le 1/y]} \quad \text{(since } X \text{ is a continuous RV)} \\ &= \frac{1 - e^{-\lambda \min(x,1/y)}}{1 - e^{-\lambda/y}}. \end{aligned}$$

Thus the conditional pdf is the derivative of this CDF with respect to x. Taking this derivative, we obtain

$$\begin{aligned} f_X(x|A, Y = y) &= \frac{d}{dx} \left(\frac{1 - e^{-\lambda \min(x,1/y)}}{1 - e^{-\lambda/y}} \right) \\ &= \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda/y}} \quad \text{for } \ 0 < x \le 1/y. \end{aligned}$$

Elsewhere $f_X(x|A.Y = y) = 0$, therefore the total solution is given as

$$f_X(x|A, Y = y) = \begin{cases} \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda/y}}, & 0 < x \le 1/y \\ 0, & \text{else.} \end{cases}$$

(c) Let $\widehat{X}$ denote the minimum mean-square error (MSE) estimate of X. Then

$$\begin{aligned}
\widehat{X} &= E[X|A, Y = y] \\
&= \int_0^\infty x f_X(x|A, Y = y) dx \\
&= \int_0^{1/y} x \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda/y}} dx \\
&= \frac{1}{\lambda} - \frac{e^{-\lambda/y}}{y\left(1 - e^{-\lambda/y}\right)}, \qquad \text{for } \ y > 0.
\end{aligned}$$

9. For two RVs of finite variance A and B, the Schwarz inequality can be written as $|E[AB^*]| \leq \sqrt{E[|A|^2]E[|B|^2]}$. So, let $A \triangleq X[n+m]$ and $B \triangleq X[n]$ to obtain $R_X[m] = E[X[n+m]X^*[n]] = E[AB^*]$, and also $E[|A|^2] = E[|B|^2] = R_X[0]$ by the WSS property of $X[n]$. Thus by Schwarz,

$$\begin{aligned}
|R_X[m]| &\leq \sqrt{R_X[0]R_X[0]} \\
&= R_X[0],
\end{aligned}$$

where we have used the fact that $R_X[0]$ is real-valued and non-negative.

10. We are given

$$X_i = \frac{2}{5}(X_{i-1} + X_{i+1}) + W_i \quad \text{for } 2 \leq i \leq 9,$$

plus $X_1 = \frac{1}{2}X_2 + \frac{5}{4}W_1$ and $X_{10} = \frac{1}{2}X_9 + \frac{5}{4}W_{10}$. Thus

$$\begin{aligned}
X_2 &= \frac{2}{5}\left(\frac{1}{2}X_2 + \frac{5}{4}W_1 + X_3\right) + W_2 \\
&= \frac{1}{5}X_2 + \frac{1}{2}W_1 + \frac{2}{5}X_3 + W_2.
\end{aligned}$$

Solving for X_2, we obtain

$$X_2 = \frac{5}{4}W_2 + \frac{5}{8}W_1 + \frac{1}{2}X_3. \tag{3}$$

Next, we take the X_3 equation, and use this result to eliminate X_2 obtaining

$$X_3 = \frac{5}{4}W_3 + \frac{5}{8}W_2 + \frac{5}{16}W_1 + \frac{1}{2}X_4. \tag{4}$$

We continue in this manner till $i = 9$, and there obtain

$$X_9 = \frac{5}{4}W_9 + \frac{5}{8}W_8 + \frac{5}{16}W_7 + \cdots + \frac{5}{1024}W_1 + \frac{1}{2}X_{10}.$$

We now solve this equation together with $X_{10} = \frac{1}{2}X_9 + \frac{5}{4}W_{10}$ to finally obtain for X_9

$$X_9 = \frac{5}{6}W_{10} + \frac{5}{3}W_9 + \frac{5}{6}W_8 + \cdots + \frac{5}{768}W_1.$$

At this point, we can work backwards, eliminating the X_{i+1} term in such equations as (3-4). At this point, we may see the general form for this answer is

$$X_i = \frac{5}{3}\sum_{k=1}^{10} \rho^{|i-k|} W_k\,, \qquad 1 \leq i \leq 10, \tag{5}$$

which can then be used for answering (a-c). Alternatively, and in terms of vectors and matrices, upon setting $\mathbf{W} = \mathbf{TX}$, we can see from the hint given in this problem, that $\mathbf{T} = \beta^2 \mathbf{A}^{-1}$ with $\beta = 3/5$, $\alpha = 2/5$, and $\rho = 1/2$. Then using the hint, we have

$$\begin{aligned}
\mathbf{X} &= \mathbf{T}^{-1}\mathbf{W} \\
&= \left(\beta^2 \mathbf{A}^{-1}\right)^{-1} \mathbf{W} \\
&= \frac{1}{\beta^2}\mathbf{AW} \\
&= \frac{1+\rho^2}{1-\rho^2}\mathbf{AW} \\
&= \frac{5}{3}\mathbf{AW},
\end{aligned}$$

which is the matrix version of (5). Using the result (5), we now answer the questions.

(a) $E[X_i] = \frac{5}{3}\sum_{k=1}^{10} \rho^{|i-k|} E[W_k] = \sum 0 = 0$, since the W_k are zero mean. Thus the mean $\boldsymbol{\mu}_X = \mathbf{0}$.

(b) Since the mean is zero, the covariance and correlation matrices agree, and so

$$\begin{aligned}
(\mathbf{K_X})_{i,j} &= E[X_i X_j] \\
&= \left(\frac{5}{3}\right)^2 E\left[\sum_{k=1}^{10} \rho^{|i-k|} W_k \sum_{l=1}^{10} \rho^{|j-l|} W_l\right] \\
&= \left(\frac{5}{3}\right)^2 \sum_{k=1}^{10}\sum_{l=1}^{10} \rho^{|i-k|}\rho^{|j-l|} E[W_k W_l] \\
&= \left(\frac{5}{3}\right)^2 \sigma^2 \sum_{k=1}^{10} \rho^{|i-k|}\rho^{|j-l|}\sigma^2 \delta_{k-l}, \qquad \text{where } \delta \text{ is the Kronecker delta,} \\
&= \left(\frac{5}{3}\right)^2 \sigma^2 \sum_{k=1}^{10}\sum_{l=1}^{10} \rho^{|i-k|}\rho^{|k-j|},
\end{aligned}$$

which can be given in matrix form as

$$\mathbf{K_X} = \left(\frac{5}{3}\right)^2 \sigma^2 \mathbf{AA}^T.$$

(c) Since the W_i are i.i.d. Laplacian with variance σ^2, we can write

$$\begin{aligned}
f_{\mathbf{W}}(\mathbf{w}) &= \prod_{i=1}^{10} f_{W_i}(w_i) \\
&= \left(\frac{1}{\sqrt{2}\sigma}\right)^{10} \exp\left(-\frac{\sqrt{2}}{\sigma}\sum_{i=1}^{10} |w_i|\right) \\
&= \left(\frac{1}{\sqrt{2}\sigma}\right)^{10} \exp\left(-\frac{\sqrt{2}}{\sigma}\mathbf{1}^T |\mathbf{w}|\right), \qquad \text{with } \mathbf{1} \text{ a vector of 1's.}
\end{aligned}$$

Now $\mathbf{W} = \mathbf{TX}$ and $\mathbf{X} = \mathbf{T}^{-1}\mathbf{W}$, so with $J = \det(\mathbf{T}^{-1}) = 1/\det \mathbf{T}$, we get

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\frac{1}{\sqrt{2}\sigma|J|}\right)^{10} \exp\left(-\frac{\sqrt{2}}{\sigma}\mathbf{1}^T |\mathbf{Tx}|\right).$$

Here $\mathbf{T}$ is given as

$$\begin{aligned}
\mathbf{T} &= \frac{3}{5}\mathbf{A}^{-1} \\
&= \begin{bmatrix} 6/5 & -2/5 & 0 & \ldots & 0 \\ -2/5 & 1 & -2/5 & 0 & \ldots \\ 0 & -2/5 & 1 & \ldots & 0 \\ \ldots & 0 & \ldots & \ldots & -2/5 \\ 0 & \ldots & 0 & -2/5 & 6/5 \end{bmatrix}.
\end{aligned}$$

11. To prove the Corollary to Theorem 8.1-1, note that the sequence of events is *decreasing* here, i.e. $B_1 \supset B_2 \supset B_3 \cdots$, so equivalently the sequence of complementary sets is increasing, i.e. $B_1^c \subset B_2^c \subset B_3^c \cdots$. So, if we apply Theorem 8.1-1 to the sequence of increasing events B_n^c, upon defining $B_\infty^c \triangleq \bigcup_{n=1}^{\infty} B_n^c$, we get that

$$\lim_{n\to\infty} P[B_n^c] = P[B_\infty^c] \ .$$

So

$$\begin{aligned}
\lim_{n\to\infty} P[B_n] &= \lim_{n\to\infty} (1 - P[B_n^c]) \\
&= 1 - \lim_{n\to\infty} P[B_n^c] \\
&= 1 - P[B_\infty^c] \\
&= P[B_\infty].
\end{aligned}$$

with $B_\infty \triangleq \bigcap_{n=1}^{\infty} B_n$ for this decreasing sequence of events. Note that by the definitions of infinite unions and intersections, $\bigcap_{n=1}^{\infty} B_n$ means the set of outcomes that are in *every* B_n and $\bigcup_{n=1}^{\infty} B_n^c$ means the set of outcomes that are in *any* of the B_n^c. Thus $\bigcup_{n=1}^{\infty} B_n^c = \left(\bigcap_{n=1}^{\infty} B_n\right)^c$, and the two expressions above for B_∞ are the same.

12. We have that $S_k \triangleq X_1 + X_2 + \ldots + X_k$ where the X_i are i.i.d. as $N(0,1)$. We start by writing

$$\begin{aligned}
f_{S_n,S_m}(s_n, s_m) &= f_{S_m,S_n-S_m}(s_m, s_n - s_m), \qquad n > m \\
&= f_{S_m}(s_m)\, f_{S_n-S_m}(s_n - s_m),
\end{aligned}$$

since for $n > m$, $S_n - S_m$ and S_n must be independent. Since they are both sums of i.i.d. standard Gaussians, they are also Gaussian, with means $E[S_m] = E[S_n - S_m] = 0$ and variances $\mathrm{Var}[S_m] = m$ and $\mathrm{Var}[S_n - S_m] = n - m (> 0)$. Hence

$$f_{S_n,S_m}(s_n, s_m) = \frac{1}{2\pi\sqrt{m(n-m)}} \exp - \left(\frac{s_m^2}{2m} + \frac{(s_n - s_m)^2}{2(n-m)} \right), \quad n > m \geq 1.$$

13. No, they need not be continuous from the left. For example, consider a discrete random variable X with the following PMF

$$P_X(x) = \begin{cases} \frac{1}{6}, & 1,2,3,4,5,6, \\ 0, & \text{else.} \end{cases}.$$

Let $F_X(x)$ be the corresponding CDF. Then, for example, $F_X(5) = \frac{5}{6}$, but for any $0 < \epsilon < 1$, we have $F_X(5-\epsilon) = \frac{4}{6}$. Therefore, $\lim_{\epsilon \longrightarrow 0} F_X(5-\epsilon) \neq F_X(5)$. Thus CDFs need not be continuous from the left.

14. (a) Denoting the outcomes as ζ_i, we have

$$\begin{aligned} \mu_X[n] &\triangleq E[X[n]] \\ &= \sum_{\zeta_i} P[\{\zeta_i\}]X[n,\zeta_i] \\ &= \frac{1}{3}\left(3\delta[n] + u[n-1] + \cos\frac{\pi n}{2}\right). \end{aligned}$$

(b)

$$\begin{aligned} R_X[m,n] &\triangleq E[X[m]X^*[n]] \\ &= \sum_{\zeta_i} P[\{\zeta_i\}]X[m,\zeta_i]X^*[n,\zeta_i] \\ &= \frac{1}{3}\left(9\delta[m]\delta[n] + u[m-1]u[n-1] + \cos\frac{\pi m}{2}\cos\frac{\pi n}{2}\right). \end{aligned}$$

(c) We can summarize the RVs $X[0]$ and $X[1]$ with the following table.

ζ_i	p	$X[0]$	$X[1]$
a	$\frac{1}{3}$	3	0
b	$\frac{1}{3}$	1	1
c	$\frac{1}{3}$	0	0

Thus $P[X[0]=3, X[1]=0] = P[\{a\}] = \frac{1}{3}$. The respective marginal probabilities are found as $P[X[0]=3] = \frac{1}{3}$ and $P[X[1]=0] = \frac{2}{3}$. Multiplying, we find

$$\begin{aligned} P[X[0] &= 3, X[1]=0] = \frac{1}{3} \\ &\neq \frac{1}{3}\frac{2}{3} \\ &= P[X[0]=3]P[X[1]=0], \end{aligned}$$

therefore the RVs $X[0]$ and $X[1]$ are not independent.

15. (a) The random variables $X[n]$ and $X[n-1]$ are jointly Gaussian distributed with zero means and covariance matrix

$$\mathbf{K} = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \quad \text{with} \quad |\rho| < 1.$$

The determinant of this matrix is $\det \mathbf{K} = \sigma^4(1-\rho^2)$, and the inverse matrix is found as

$$\mathbf{K}^{-1} = \frac{1}{\sigma^4(1-\rho^2)}\begin{bmatrix} \sigma^2 & -\rho\sigma^2 \\ -\rho\sigma^2 & \sigma^2 \end{bmatrix}.$$

We can then write their joint pdf as

$$f_X(x_n, x_{n-1}) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2\sigma^2(1-\rho^2)}\left(x_n^2 - 2\rho x_n x_{n-1} + x_{n-1}^2\right)\right).$$

Also the marginal pdf for $X[n-1]$ is given directly as

$$f_X(x_{n-1}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x_{n-1}^2}{2\sigma^2}\right).$$

We can the write the conditional density

$$\begin{aligned} f_X(x_n|x_{n-1}) &= \frac{f_X(x_n, x_{n-1})}{f_X(x_{n-1})} \\ &= \frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2\sigma^2(1-\rho^2)}\left(x_n - \rho x_{n-1}\right)^2\right), \end{aligned}$$

after recognizing the perfect square $x_n^2 - 2\rho x_n x_{n-1} + \rho^2 x_{n-1}^2 \equiv (x_n - \rho x_{n-1})^2$. Recognizing that this conditional density is $N(\rho x_{n-1}, \sigma^2(1-\rho^2))$, we can immediately write its conditional mean

$$E[X[n]|X[n-1]] = \rho X[n-1].$$

(b) This predictor minimizes the mean square error over all functions $gX[n-1))$, i.e. it minimizes $E[(X[n] - g(X[n-1])^2]$ over all functions g. c.f. Example 4.3-4.

16. (a)

$$\begin{aligned} \mu_Y[n] &\triangleq E[Y[n]] \\ &= E\left[\sum_k h[k]X[n-k]\right] \\ &= \sum_k h[k]\mu_X[n-k] \\ &= \frac{1}{2}\mu_X[n] + \frac{1}{2}\mu_X[n-1]. \end{aligned}$$

(b)

$$\begin{aligned} R_{YY}[n_1, n_2] &\triangleq E[Y[n_1]Y^*[n_2]] \\ &= \sum_{k,l} h[k]h^*[l]R_{XX}[n_1-k, n_2-l] \\ &= \frac{1}{4}\left(R_{XX}[n_1, n_2] + R_{XX}[n_1-1, n_2] + R_{XX}[n_1, n_2-1] + R_{XX}[n_1-1, n_2-1]\right). \end{aligned}$$

(c)

$$\begin{aligned} K_{YY}[n_1, n_2] &\triangleq E[(Y[n_1] - \mu_Y[n_1])(Y^*[n_2] - \mu_Y^*[n_2])] \\ &= R_{YY}[n_1, n_2] - \mu_Y[n_1]\mu_Y^*[n_2] \\ &= \frac{1}{4}\left(R_{XX}[n_1, n_2] + R_{XX}[n_1-1, n_2] + R_{XX}[n_1, n_2-1] + R_{XX}[n_1-1, n_2-1]\right) \\ &\quad - \left(\frac{1}{2}\mu_X[n_1] + \frac{1}{2}\mu_X[n_1-1]\right)\left(\frac{1}{2}\mu_X^*[n_2] + \frac{1}{2}\mu_X^*[n_2-1]\right) \\ &= \frac{1}{4}\left(K_{XX}[n_1, n_2] + K_{XX}[n_1-1, n_2] + K_{XX}[n_1, n_2-1] + K_{XX}[n_1-1, n_2-1]\right). \end{aligned}$$

(d) Set $\mathbf{Y} \triangleq \begin{bmatrix} Y[n_1] \\ Y[n_2] \end{bmatrix}$, then $\mathbf{Y}$ is distributed as $N(\boldsymbol{\mu}_Y, \mathbf{K}_{YY})$, where

$$\boldsymbol{\mu}_Y = \begin{bmatrix} \mu_Y[n_1] \\ \mu_Y[n_2] \end{bmatrix} \quad \text{and} \quad \mathbf{K}_{YY} = \begin{bmatrix} K_{YY}[n_1, n_1] & K_{YY}[n_1, n_2] \\ K_{YY}[n_2, n_1] & K_{YY}[n_2, n_2] \end{bmatrix}.$$

The pdf of vector $\mathbf{Y}$ is given as

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)|\det \mathbf{K}_{YY}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}_Y)^T \mathbf{K}_{YY}^{-1}(\mathbf{y} - \boldsymbol{\mu}_Y)\right).$$

17. We need the joint pdf $f_T(t_2, t_1; 10, 5)$. Now

$$\begin{aligned} T[10] &= \sum_{k=1}^{10} \tau[k] \\ &= T[5] + \sum_{k=6}^{10} \tau[k]. \end{aligned}$$

Calling $X \triangleq \sum_{k=6}^{10} \tau[k]$, we see that by definition, X and $T[5]$ are independent. This since $T[5]$ is a sum of earlier $\tau[k]'s$ not included in the sum that is X. Thus

$$\begin{aligned} f_T(t_2, t_1; 10, 5) &= f_T(t_2|t_1; 10, 5) f_T(t_1; 5) \\ &= f_T(t_2 - t_1; (10 - 5)) f_T(t_1; 5) \\ &= f_T(t_2 - t_1; 5) f_T(t_1; 5) \\ &= \frac{(\lambda(t_2 - t_1))^4}{4!} \lambda e^{-\lambda(t_2 - t_1)} \frac{(\lambda t_1)^4}{4!} \lambda e^{-\lambda t_1}, \quad t_2 \geq t_1 \geq 0. \end{aligned}$$

18. (a)

$$T[n] = \sum_{k=1}^{n} \tau[k], \qquad n \geq 1.$$

$$\begin{aligned} \Phi_T(\omega; n) &\triangleq E[e^{j\omega T[n]}] \\ &= \prod_{k=1}^{n} E\left[e^{+j\omega\tau[k]}\right] \\ &= (\Phi_\tau(\omega))^n. \end{aligned}$$

In turn

$$\begin{aligned} \Phi_\tau(\omega) &= E[e^{j\omega\tau}] \\ &= \int_0^\infty \lambda e^{-\lambda\tau} e^{+j\omega\tau} d\tau \\ &= \lambda \left(\left. \frac{e^{(-\lambda + j\omega)\tau}}{-\lambda + j\omega} \right|_0^\infty \right) \\ &= \frac{\lambda}{\lambda - j\omega}, \quad \text{since} \quad \lambda > 0. \end{aligned}$$

Thus

$$\Phi_T(\omega; n) = \left(\frac{\lambda}{\lambda - j\omega}\right)^n .$$

(b)

$$\begin{aligned}
\mu_T[n] &= m_1 \\
&= \frac{1}{j}\Phi_T^{(1)}(0; n) \\
&= \frac{1}{j}\lambda^n \left(\left. \frac{(\lambda - j\omega)^n \cdot 0 + nj\,(\lambda - j\omega)^{n-1}}{(\lambda - j\omega)^{2n}} \right|_{\omega=0} \right) \\
&= \frac{1}{j}\lambda^n \frac{nj}{\lambda^{n+1}} \\
&= \frac{n}{\lambda}, \qquad n \geq 1.
\end{aligned}$$

This answer is correct because each of the n interarrival times $\tau[k]$ has average value $1/\lambda$.

19. No. Since $\tau[n]$ and $\tau[n-1]$ can have any joint pdf. which has not been specified. If $\tau[n]$ has independent increments and if the increments $\tau[n] - \tau[n-1]$ are identically distributed, then because of $T[n] = \sum_{k=1}^{n} \tau[k]$, it follows that

$$\Phi_T(\omega) = \Phi_\tau^n(\omega), \text{ which implies } \quad \Phi_\tau(\omega) = \sqrt[n]{\Phi_T(\omega)},$$

where we take the positive real nth root since it must hold that $\Phi_\tau(0) = 1$. By this result, we have that $\tau[n]$ is exponentially distributed with parameter λ. In this case, the joint pdf would be

$$\begin{aligned}
f_T(t_n, t_{n-1}; n, n-1) &= f_T(t_{n-1}; n-1) f_\tau(t_n - t_{n-1}; n) \\
&= \frac{(\lambda t_{n-1})^{n-2}}{(n-2)!} \lambda e^{-\lambda t_{n-1}} \cdot \lambda e^{-\lambda(t_n - t_{n-1})} u(t_n - t_{n-1}) u(t_{n-1}).
\end{aligned}$$

Actually, the following weaker assumption will do here: that $T[n-1]$ and $T[n] - T[n-1]$ are independent for all $n \geq 1$.

20. We have $X[n, \boldsymbol{\zeta}] = \sum_{i=-\infty}^{i=+\infty} A(\zeta_i) h[n-i]$, with $h[n] = \begin{cases} 1/4 & n = -1 \\ 1/2 & n = 0 \\ 1/4 & n = +1 \\ 0 & \text{else,} \end{cases}$, thus

(a) $E[X[n]] = \sum_{i=-1}^{i=+1} E[A_i] h[n-i] = \lambda(\frac{1}{4} + \frac{1}{2} + \frac{1}{4}) = \lambda = \mu$. Here, for the vector outcome $\boldsymbol{\zeta}$, with ζ_i as it's ith component, we call the RV $A(\zeta_i) \triangleq A_i$.

(b) $\sigma_X^2(n) = E[(X[n]-\lambda)^2] = E[X^2[n]] - \lambda^2$, with

$$\begin{aligned}
E[X^2[n]] &= \sum_{i=n-1}^{i=n+1}\sum_{j=n-1}^{j=n+1} E[A_iA_j]h[n-i]h[n-j] \\
&= \sum_{i=n-1}^{i=n+1} E[A_i^2]h^2[n-i] + \sum_{i\neq j} E[A_i]E[A_j]h[n-i]h[n-j] \\
&= (\lambda^2+\lambda)\left(\frac{1}{16}+\frac{1}{4}+\frac{1}{16}\right) + \lambda^2\left(\frac{1}{4}\frac{1}{2}+\frac{1}{4}\frac{1}{4}+\frac{1}{2}\frac{1}{4}+\frac{1}{4}\frac{1}{2}+\frac{1}{4}\frac{1}{4}\right) \\
&= \frac{3}{8}(\lambda^2+\lambda) + \frac{5}{8}\lambda^2 = \frac{3}{8}\lambda + \lambda^2.
\end{aligned}$$

Thus $\sigma_X^2[n] = \frac{3}{8}\lambda + \lambda^2 - \lambda^2 = \frac{3}{8}\lambda = \sigma^2$ and $X[n] : N(\lambda, 3\lambda/8)$.

(c) Since the $X's$ will be correlated, we need to specify the correlation coefficient ρ to complete the expression for this joint pdf.

$$\begin{aligned}
E[X[n]X[n+1]] &= E\left[\sum_{i=n-1}^{i=n+1} A_ih[n-i]\sum_{j=n}^{j=n+2} A_jh[n+1-j]\right] \\
&= \sum_{i=n-1}^{i=n+1}\sum_{j=n}^{j=n+2} E[A_iA_j]h[n-i]h[n+1-j] \\
= (\lambda^2+\lambda)\left(\frac{1}{8}+\frac{1}{8}\right) &+ \lambda^2\left(\frac{1}{4}\frac{1}{4}+\frac{1}{4}\frac{1}{2}+\frac{1}{4}\frac{1}{4}+\frac{1}{2}\frac{1}{2}+\frac{1}{2}\frac{1}{4}+\frac{1}{4}\frac{1}{4}+\frac{1}{4}\frac{1}{4}\right) \\
&= \frac{1}{4}\lambda^2 + \frac{1}{4}\lambda + \frac{3}{4}\lambda^2 \\
&= \frac{1}{4}\lambda + \lambda^2.
\end{aligned}$$

Then

$$\begin{aligned}
\rho &= \frac{\mathrm{cov}[X[n]X[n+1]]}{\sigma_X[n]\sigma_X[n+1]} \\
&= \frac{E[X[n]X[n+1]] - \lambda^2}{\sigma^2} \\
&= \frac{2}{3}.
\end{aligned}$$

So, the joint pdf becomes

$$\begin{aligned}
f_{X[n],X[n+1]}(x_1,x_2) &= \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu)^2}{\sigma^2} - 2\rho\frac{(x_1-\mu)(x_2-\mu)}{\sigma^2}\frac{(x_2-\mu)^2}{\sigma^2}\right]\right\} \\
&= \frac{1}{2\pi\frac{3}{8}\lambda\sqrt{1-\left(\frac{2}{3}\right)^2}}\exp\left\{-\frac{1}{2(1-\left(\frac{2}{3}\right)^2)}\left[\frac{(x_1-\lambda)^2}{\frac{3}{8}\lambda} - 2\frac{2}{3}\frac{(x_1-\lambda)(x_2-\lambda)}{\frac{3}{8}\lambda}\frac{(x_2-\lambda)^2}{\frac{3}{8}\lambda}\right]\right\} \\
&= \frac{1}{\frac{\sqrt{5}}{4}\pi\lambda}\exp\left\{-\frac{24}{10}\left[\frac{(x_1-\lambda)^2}{\lambda} - \frac{4}{3}\frac{(x_1-\lambda)(x_2-\lambda)}{\lambda}\frac{(x_2-\lambda)^2}{\lambda}\right]\right\}.
\end{aligned}$$

(Alternate simple solution to parts a and b using CFs)

$$\begin{aligned}\Phi_X(\omega) &= E[e^{j\omega X}] \\ &= E\left[e^{+j\omega(\frac{1}{4}A_{-1}+\frac{1}{2}A_0+\frac{1}{4}A_1)}\right] \\ &= \Phi_A^2(\frac{\omega}{4})\Phi_A(\frac{\omega}{2}).\end{aligned}$$

Now $\Phi_A(\omega) = E[e^{+j\omega A}] = \exp(j\omega\lambda - \frac{1}{2}\omega^2\lambda)$ since $A : N(\lambda, \lambda)$. Thus

$$\Phi_A(\frac{\omega}{4}) = \exp(j\omega\frac{\lambda}{4} - \frac{1}{2}\omega^2\frac{\lambda}{16}) \text{ and } \Phi_A(\frac{\omega}{2}) = \exp(j\omega\frac{\lambda}{2} - \frac{1}{2}\omega^2\frac{\lambda}{4}),$$

so

$$\Phi_X(\omega) = \exp\left(j\omega\lambda - \frac{1}{2}\omega^2\frac{3}{8}\lambda\right) \quad \text{and so} \quad X[n] : N(\lambda, 3\lambda/8).$$

21. First we compute

$$\begin{aligned}f_X(x_1) &= \int_{-\infty}^{+\infty} f_X(x_1|x_0)\delta(x_0)dx_0 \\ &= \alpha e^{-\alpha x_1}u(x_1).\end{aligned}$$

Then

$$\begin{aligned}f_X(x_2) &= \int_{-\infty}^{+\infty} f_X(x_2|x_1)f_X(x_1)dx_1 \\ &= \int_{-\infty}^{+\infty} \alpha e^{-\alpha(x_2-x_1)}u(x_2-x_1)\alpha e^{-\alpha x_1}u(x_1)dx_1 \\ &= \alpha^2 e^{-\alpha x_2}\int_{-\infty}^{+\infty} u(x_2-x_1)u(x_1)dx_1 \\ &= \alpha^2 e^{-\alpha x_2}u(x_2)\left(\int_0^{x_2} dx_1\right) \\ &= \alpha^2 x_2 e^{-\alpha x_2}u(x_2).\end{aligned}$$

(b) Doing this again for $n = 3$, we would get $f_X(x_3) = \frac{\alpha^3 x_3^2}{2!}e^{-\alpha x_3}u(x_3)$, thus we guess the Erlang pdf

$$f_X(x_n) = \frac{\alpha^n x_n^{n-1}}{(n-1)!}e^{-\alpha x_n}u(x_n).$$

Then, using mathematical induction, we must calculate

$$\begin{aligned}
f_X(x_n) &= \int_{-\infty}^{+\infty} f_X(x_n|x_{n-1})f_X(x_{n-1})dx_{n-1} \\
&= \int_{-\infty}^{+\infty} \alpha e^{-\alpha(x_n-x_{n-1})}u(x_n-x_{n-1})f_X(x_{n-1})dx_{n-1} \\
&= \alpha e^{-\alpha x_n}\left(\int_0^{x_n} e^{\alpha x_{n-1}} f_X(x_{n-1})dx_{n-1}\right)u(x_n) \\
&= \alpha e^{-\alpha x_n}\left(\int_0^{x_n} e^{\alpha x_{n-1}} \frac{\alpha^{n-1}x_{n-1}^{n-2}}{(n-2)!}e^{-\alpha x_{n-1}}dx_{n-1}\right)u(x_n) \\
&= \frac{\alpha^n}{(n-2)!}e^{-\alpha x_n}\left(\int_0^{x_n} x_{n-1}^{n-2}dx_{n-1}\right)u(x_n) \\
&= \frac{\alpha^n}{(n-2)!}e^{-\alpha x_n}\frac{x_n^{n-1}}{n-1}u(x_n) \\
&= \frac{\alpha^n x_n^{n-1}}{(n-1)!}e^{-\alpha x_n}u(x_n), \qquad \text{as was to be shown.}
\end{aligned}$$

22. Let the system be represented by operator L as $y[n] = L\{x[n]\}$. From the definition$h[n] = L\{\delta[n]\}$ with $\delta[n]$ being the discrete time impulse function $\delta[n] \triangleq \begin{cases} 1, & n = 0, \\ 0, & \text{else.} \end{cases}$ Next, using the shifting representation, we write the input sequence as $x[n] = \sum_{k=-\infty}^{+\infty} x[k]\delta[n-k]$. Then we can compute

$$\begin{aligned}
y[n] &= L\{x[n]\} \\
&= L\left\{\sum_{k=-\infty}^{+\infty} x[k]\delta[n-k]\right\} \\
&= \sum_{k=-\infty}^{+\infty} x[k]L\{\delta[n-k]\}, \qquad \text{by linearity for a continuous operator } L, \\
&= \sum_{k=-\infty}^{+\infty} x[k]h[n-k]. \\
&= x[n] * h[n].
\end{aligned}$$

Therefore $Y[n] = X[n] * h[n]$ too. Note that in order to interchange the operator L and the infinite summation operator $\sum_{k=-\infty}^{+\infty}$, we generally need that $h[n]$ be absolutely summable, i.e. $\sum_{n=-\infty}^{+\infty} |h[n]| < \infty$, a stable system. Stable operators L are *continuous* in the sense that a small change in the input sequence x results in a bounded change in the output sequence y.

(b) $A(\omega) \triangleq \sum_{n=-\infty}^{+\infty} a[n]e^{-j\omega n}$ and so,

$$\begin{aligned}
a[n] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} A(\omega)e^{+j\omega n}d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{n=-\infty}^{+\infty} a[m]e^{-j\omega m}\right)e^{+j\omega n}d\omega \\
&= \sum_{n=-\infty}^{+\infty} a[m]\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{+j\omega(n-m)}d\omega\right), \quad \text{by interchanging the infinite sum and the integral,} \\
&= \sum_{n=-\infty}^{+\infty} a[m]\delta[n-m] \\
&= a[n],
\end{aligned}$$

where the interchange of the infinite sum and the integral is permitted if the sequence a is absolutely summable, i.e. $\sum_{n=-\infty}^{+\infty} |a[n]| < \infty$.

(c) We have $y[n] = \sum_{k=-\infty}^{+\infty} x[k]h[n-k]$, thus

$$\begin{aligned}
Y(\omega) &= \sum_{n=-\infty}^{+\infty}\left(\sum_{k=-\infty}^{+\infty} x[k]h[n-k]\right)e^{-j\omega n} \\
&= \sum_{k=-\infty}^{+\infty} x[k]\left(\sum_{n=-\infty}^{+\infty} h[n-k]e^{-j\omega n}\right), \quad \text{by interchanging the infinite sums,} \\
&= \sum_{k=-\infty}^{+\infty} x[k]e^{-j\omega k}\left(\sum_{n=-\infty}^{+\infty} h[n-k]e^{-j\omega n}e^{+j\omega k}\right) \\
&= \sum_{k=-\infty}^{+\infty} x[k]e^{-j\omega k}\left(\sum_{n=-\infty}^{+\infty} h[n-k]e^{-j\omega(n-k)}\right) \\
&= \sum_{k=-\infty}^{+\infty} x[k]e^{-j\omega k}\left(H(\omega)\right) \\
&= H(\omega)\sum_{k=-\infty}^{+\infty} x[k]e^{-j\omega k} \\
&= H(\omega)X(\omega).
\end{aligned}$$

Note that the interchange of the infinite sums in the steps above can be justified if the infinite sum $\sum_{k=-\infty}^{+\infty} x[k]h[n-k]$ converges uniformly. This occurs when $\sum_{k=-\infty}^{+\infty} |x[k]|\cdot|h[n-k]| < \infty$.

23. (a) Using Z transforms

$$Y(z) + \alpha z^{-1}Y(z) = X(z) \quad \text{and} \quad X(z) = \frac{1}{1-\beta z^{-1}},$$

so

$$Y(z) = \frac{1}{(1+\alpha z^{-1})(1-\beta z^{-1})} = \frac{z^2}{(z+\alpha)(z-\beta)}.$$

Now via the residue method, evaluating the complex integral

$$y[n] = \frac{1}{2\pi j} \oint Y(z) z^{n-1} dz, \qquad \text{see Appendix A.3 (discrete time),}$$

$$\begin{aligned} y[n] &= \frac{(-\alpha)^{n+1}}{-\alpha-\beta} + \frac{\beta^{n+1}}{\alpha+\beta} \\ &= \frac{1}{\alpha+\beta}\left(\beta^{n+1} - (-\alpha)^{n+1}\right) \quad n \geq 0. \end{aligned}$$

The answer is zero for $n < 0$, so the full answer is

$$y[n] = \begin{cases} \frac{1}{\alpha+\beta}\left(\beta^{n+1} - (-\alpha)^{n+1}\right), & n \geq 0, \\ 0, & n < 0. \end{cases}$$

Alternatively, we can use the partial fraction method, and first express

$$\begin{aligned} Y(z) &= \frac{A}{1+\alpha z^{-1}} + \frac{B}{1-\beta z^{-1}} \\ &= \frac{\alpha/(\alpha+\beta)}{1+\alpha z^{-1}} + \frac{\beta/(\alpha+\beta)}{1-\beta z^{-1}}, \end{aligned}$$

and upon inverse Z-transform obtaining the same answer.

24. (a) We are given that ρ is a real constant, but in general α could be complex.

$$\begin{aligned} K_{YY}[m] &\triangleq E[Y[n+m]Y^*[n]] \\ &= E\left[(X[n+m] - \alpha X[n+m-1])(X[n] - \alpha X[n-1])^*\right] \\ &= K_{XX}[m] - \alpha K_{XX}[m-1] - \alpha^* K_{XX}[m+1] + |\alpha|^2 K_{XX}[m] \\ &= (1+|\alpha|^2)K_{XX}[m] - \alpha K_{XX}[m-1] - \alpha^* K_{XX}[m+1] \\ &= \sigma^2\left[(1+|\alpha|^2)\rho^{|m|} - \alpha\rho^{|m-1|} - \alpha^*\rho^{|m+1|}\right]. \end{aligned}$$

(b) To get white noise, we try α real and take $m \geq 1$, then we set

$$\begin{aligned} 0 &= (1+\alpha^2)\rho^m - \alpha\rho^{m-1} - \alpha\rho^{m+1} \\ &= \rho^m(1+\alpha^2 - \alpha/\rho - \alpha\rho) \\ &\Longrightarrow \alpha = \rho. \end{aligned}$$

This also works, i.e. gives zero for $K_{YY}[m]$ for $m < 0$, thus $\alpha = \rho$ is a solution. The value $\alpha = \rho^{-1}$ also works to produce white noise at the system output.

(c) For $m = 0$, we then get the variance of the white noise sequence Y

$$\begin{aligned} \sigma_Y^2 &= \sigma^2(1+\alpha^2 - \alpha\rho - \alpha\rho) \\ &= \sigma^2(1-\rho^2), \quad \text{with the choice } \alpha = \rho. \end{aligned}$$

Alternatively, with the choice $\alpha = \rho^{-1}$, we get $\sigma_Y^2 = (\rho^{-2} - 1)$.

25. (a) From the given LCCDE, we can seethat $X[n-1]$ is a linear combination of only the $W[n-i]'$ for $i > 0$. Thus, since the $W[n]$ are all jointly independent, it follows that $X[n-1]$ and $W[n]$ are independent.

(b)

$$\begin{aligned}
\Phi_{X[n]}(\omega) &\triangleq E[e^{+j\omega X[n]}] \\
&= E[e^{+j\omega(\rho X[n-1]+W[n])}] \\
&= E[e^{+j\omega\rho X[n-1]}]E[e^{+j\omega W[n]}], \quad \text{by independence,} \\
&= \Phi_{X[n-1]}(\rho\omega)\Phi_W(\omega) \\
&= \Phi_X(\rho\omega)\Phi_W(\omega), \qquad \text{since } \; X[n] \text{ is stationary.}
\end{aligned}$$

Note that we must have $|\rho| < 1$ for stationarity.

(c) Since $W[n]$ is Gaussian with zero mean, we have

$$\Phi_W(\omega) = e^{-\frac{1}{2}\sigma_W^2\omega^2}.$$

Now from the answer to (b), we have upon iteration,

$$\begin{aligned}
\Phi_X(\omega) &= \Phi_X(\rho\omega)\Phi_W(\omega) \\
&= \Phi_X(\rho^2\omega)\Phi_W(\rho\omega)\Phi_W(\omega) \\
&= \Phi_X(\rho^3\omega)\Phi_W(\rho^2\omega)\Phi_W(\rho\omega)\Phi_W(\omega) \\
&\cdots \\
&= \Phi_X(\rho^k\omega)\Phi_W(\rho^{k-1}\omega)\Phi_W(\rho^{k-2}\omega)\cdots\Phi_W(\omega).
\end{aligned}$$

Now in the limit, as $k \to \infty$, the condition $|\rho| < 1$ forces $\rho^k\omega$ to zero for any finite ω, therefore assuming continuity in the CF, we get

$$\begin{aligned}
\Phi_X(\omega) &= \prod_{k=0}^{\infty}\Phi_W(\rho^k\omega) \\
&= \exp\left(-\sum_{k=0}^{\infty}\frac{1}{2}\sigma_W^2\rho^{2k}\omega^2\right) \\
&= \exp\left(-\frac{1}{2}\sigma_W^2\omega^2\sum_{k=0}^{\infty}\rho^{2k}\right) \\
&= \exp\left(-\frac{1}{2}\sigma_W^2\omega^2\frac{1}{1-\rho^2}\right) \\
&= \exp\left(-\frac{1}{2}\frac{\sigma_W^2}{1-\rho^2}\omega^2\right).
\end{aligned}$$

(d) We recognize the CF of Gaussian noise in the result of (c), therefore it must be that $\sigma_X^2 = \frac{\sigma_W^2}{1-\rho^2}$.

26. (a) From the problem, $Y[n] = h[n] * [W[n] + X[n]]$, so

$$\begin{aligned}
\mu_Y[n] &= h[n] * (\mu_W[n] + 3) \\
&= \sum_{k=0}^{\infty} \rho^k (\mu_W[n-k] + 3) \\
&= \sum_{k=0}^{\infty} \rho^k (2 + 3) \\
&= 5 \sum_{k=0}^{\infty} \rho^k \\
&= \frac{5}{1 - \rho.}
\end{aligned}$$

(b) The second moment of the real-valued random sequence Y is given as:

$$\begin{aligned}
E[Y^2[n]] &= E\left[\left(\sum_{k=0}^{\infty} h[k](W[n-k] + 3)\right)^2\right] \\
&= \sum_{(k,l)\geq 0}^{\infty} h[k]h[l]E[(W[n-k] + 3)\,(W[n-l] + 3)] \\
&= \sum_{(k,l)\geq 0}^{\infty} h[k]h[l](\sigma_W^2 \delta[l-k] + 4 + 9 + 6 + 6) \\
&= \sum_{(k,l)\geq 0}^{\infty} h[k]h[l](\sigma_W^2 \delta[l-k] + 25) \\
&= \sum_{k=0}^{\infty} h^2[k]\sigma_W^2 + \sum_{(k,l)\geq 0}^{\infty} h[k]h[l](25) \\
&= \left(\sum_{k=0}^{\infty} h^2[k]\right) \sigma_W^2 + \left(\sum_{k=0}^{\infty} h[k]\right)^2 25 \\
&= \left(\sum_{k=0}^{\infty} \rho^{2k}\right) \sigma_W^2 + \left(\sum_{k=0}^{\infty} \rho^k\right)^2 25 \\
&= \frac{\sigma_W^2}{1 - \rho^2} + \frac{25}{(1-\rho)^2}.
\end{aligned}$$

(c) For the covariance function of Y, we have

$$\begin{aligned}
K_{YY}[m,n] &= \sum_{(k,l)\geq 0}^{\infty} h[k]h[l]K_{WW}[m-k,n-l] \\
&= \sum_{(k,l)\geq 0}^{\infty} h[k]h[l]\sigma_W^2\delta[m-k-(n-l)] \\
&= \sum_{(k,l)\geq 0}^{\infty} h[k]h[l]\sigma_W^2\delta[(m-n)-(k-l)] \\
&= \sum_{(k,l)\geq 0}^{\infty} h[k]h[l]\sigma_W^2\delta[(m-n)-(k-l)] \\
&= \sum_{k=0}^{\infty} h[k]h[k-(m-n)]\sigma_W^2 \\
&= g(m-n),
\end{aligned}$$

where $g(m) = K_{YY}[m]$, the WSS covariance function. Continuing on,

$$\begin{aligned}
K_{YY}[m] &= \sum_{k=0}^{\infty} h[k]h[k-m]\sigma_W^2 \\
&= \sum_{k=\max(0,m)}^{\infty} \rho^k\rho^{k-m}\sigma_W^2 \\
&= \left(\sum_{k=\max(0,m)}^{\infty} \rho^{2k}\right)\rho^{-m}\sigma_W^2 \\
&= \frac{\rho^{2\max(0,m)}}{1-\rho^2}\rho^{-m}\sigma_W^2 \\
&= \rho^{|m|}\frac{\sigma_W^2}{1-\rho^2}.
\end{aligned}$$

Thus $K_{YY}[m,n] = K_{YY}[m-n] = \rho^{|m-n|}\frac{\sigma_W^2}{1-\rho^2}$.

27. We will show $E[X[n]X[n+1]] \neq 0$. Assume that α is real. Then

$$X[n] = \sum_{m=1}^{n} \alpha^{n-m}W[m] \quad \text{and} \quad X[n+1] = \sum_{l=1}^{n+1} \alpha^{n+1-l}W[l].$$

So

$$
\begin{aligned}
E[X[n]X[n+1]] &= \sum_{m=1}^{n}\sum_{l=1}^{n+1} \alpha^{n-m}\alpha^{n+1-l}E[W[m]W[l]] \\
&= \alpha^{2n+1}\sum_{l=1}^{n}\alpha^{-2l} \\
&= \alpha^{2n-1}\sum_{i=0}^{n-1}\alpha^{-2i} \\
&= \alpha\left(\frac{1-\alpha^{2n}}{1-\alpha^2}\right) \\
&= \alpha \cdot \mathrm{Var}[X] \\
&\neq 0.
\end{aligned}
$$

28.

29.

30. (a) $K_{XX}[0]$ must be non-negative and here $K_{XX}[0] = p^2 - \mu^2$, so $p^2 \geq \mu^2$. Set $\sigma^2 \triangleq p^2 - \mu^2$ (≥ 0).

(b) The $N \times N$ covariance matrix is then

$$
\mathbf{K_{XX}} = \begin{bmatrix} \sigma^2 & & -\mu^2 \\ & \ddots & \\ -\mu^2 & & \sigma^2 \end{bmatrix},
$$

so for the all 1s vector $\mathbf{a}$, we get $\mathbf{aKa}^\dagger = N\sigma^2 - N(N-1)\mu^2 \geq 0$, so it must be that

$$
\begin{aligned}
\mu^2 &\leq \frac{N\sigma^2}{N(N-1)} \\
&= \frac{\sigma^2}{N-1}.
\end{aligned}
$$

(c) By taking a sequence of ever increasing N values, we conclude that μ must be zero, for any finite p^2 and hence σ^2.

31.

32. We are given $R_{XX}[m] = 10e^{-\lambda_1|m|} + 5e^{-\lambda_2|m|}$ with $\lambda_1 > 0$ and $\lambda_2 > 0$. We assume $\lambda_1 \neq \lambda_2$

and offer the general solution.

$$
\begin{aligned}
S_{XX}(\omega) &\triangleq \sum_{m=-\infty}^{+\infty} R_{XX}[m]e^{-j\omega m} \\
&= \sum_{m=-\infty}^{+\infty} 10e^{-\lambda_1|m|}e^{-j\omega m} + \sum_{m=-\infty}^{+\infty} 5e^{-\lambda_2|m|}e^{-j\omega m} \\
&= 10\left(\sum_{m=0}^{+\infty} e^{-\lambda_1 m}e^{-j\omega m} + \sum_{m=-\infty}^{-1} e^{+\lambda_1 m}e^{-j\omega m}\right) \\
&\quad +5\left(\sum_{m=0}^{+\infty} e^{-\lambda_2 m}e^{-j\omega m} + \sum_{m=-\infty}^{-1} e^{+\lambda_2 m}e^{-j\omega m}\right) \\
&= 10\left(\sum_{m=0}^{+\infty} e^{-(\lambda_1+j\omega)m} + \sum_{m=-\infty}^{0} e^{+(\lambda_1-j\omega)m} - 1\right) \\
&\quad +5\left(\sum_{m=0}^{+\infty} e^{-(\lambda_2+j\omega)m} + \sum_{m=-\infty}^{0} e^{+(\lambda_2-j\omega)m} - 1\right) \\
&= 10\left(\sum_{m=0}^{+\infty} e^{-(\lambda_1+j\omega)m} + \sum_{m'=0}^{+\infty} e^{-(\lambda_1-j\omega)m'} - 1\right) \\
&+ 5\left(\sum_{m=0}^{+\infty} e^{-(\lambda_2+j\omega)m} + \sum_{m'=0}^{+\infty} e^{-(\lambda_2-j\omega)m'} - 1\right), \quad \text{with sub } m' \triangleq -m, \\
&= 10\left(\frac{1}{1-e^{-(\lambda_1+j\omega)}} + \frac{1}{1-e^{-(\lambda_1-j\omega)}} - 1\right) + 5\left(\frac{1}{1-e^{-(\lambda_2+j\omega)}} + \frac{1}{1-e^{-(\lambda_2-j\omega)}} - 1\right) \\
&= 10\left(\frac{1-e^{-2\lambda_1}}{1-2\cos\omega\, e^{-\lambda_1}+e^{-2\lambda_1}}\right) + 5\left(\frac{1-e^{-2\lambda_2}}{1-2\cos\omega\, e^{-\lambda_2}+e^{-2\lambda_2}}\right).
\end{aligned}
$$

33.

34.

35. (a) $S_{YY}(\omega) = |H(\omega)|^2(S_{XX}(\omega) + S_{VV}(\omega))$.
(b) $Y[n] = \sum_{k=-\infty}^{+\infty} h[k](X[n-k]+V[n-k])$. Now $E[X[n+m]Y^*[n]] = \sum_{k=-\infty}^{+\infty} h^*[k]E[X[n+m]X^*[n-k]]$, so

$$
\begin{aligned}
R_{XY}[m] &= \sum_{k=-\infty}^{+\infty} h^*[k]R_{XX}[m+k] \\
&= \sum_{k'=-\infty}^{+\infty} h^*[-k']R_{XX}[m-k'], \quad \text{with} \quad k' \triangleq -k, \\
&= R_{XX}[m] * h^*[-m].
\end{aligned}
$$

Hence $S_{XY}(\omega) = H^*(\omega)S_{XX}(\omega)$.

36. For this system,

$$h[n] = \frac{1}{5}\left(\delta[n+2] + \delta[n+1] + \delta[n] + \delta[n-1] + \delta[n-2]\right)$$

and

$$\begin{aligned} H(\omega) &= \frac{1}{5}(1 + 2\cos\omega + 2\cos 2\omega) \\ &= \frac{1}{5}\frac{\sin\frac{5}{2}\omega}{\sin\frac{1}{2}\omega}. \end{aligned}$$

Then

(a)

$$\begin{aligned} S_{YY}(\omega) &= |H(\omega)|^2\, S_{XX}(\omega) \\ &= \frac{1}{25}(1 + 2\cos\omega + 2\cos 2\omega)^2 \cdot 2 \\ &= \frac{2}{25}\left(\frac{\sin\frac{5}{2}\omega}{\sin\frac{1}{2}\omega}\right)^2. \end{aligned}$$

(b)

$$\begin{aligned} R_{YY}[m] &= h[m] * h[-m] * [\delta[m] \\ &= \frac{2}{25}\mathrm{triag}[m]. \end{aligned}$$

Here, the triangular finite-support sequence triag[·] is specified as follows:

n	0	± 1	± 2	± 3	± 4	else
triag[n]	5	4	3	2	1	0

37. (a)

$$\begin{aligned} S_{WW}(\omega) &= |G(\omega)|^2 S_{YY}(\omega) \\ &= \frac{1}{S_{YY}(\omega)} S_{YY}(\omega) \\ &= 1. \end{aligned}$$

For cross-power spectral density, we go back to the time domain first, $R_{XW}[m] = g[m] * R_{XX}[m]$, so

$$\begin{aligned} S_{XW}(\omega) &= G(\omega) S_{XX}(\omega) \\ &= \frac{1}{\sqrt{S_{YY}(\omega)}} S_{XX}(\omega). \end{aligned}$$

(b) We have the FIR estimator

$$\widehat{X}[n] = \sum_{k=0}^{N-1} h[k]W[n-k],$$

with orthogonality condition

$$(\widehat{X}[n] - X[n]) \perp W[m], \text{ for } m = n, n-1, ..., n-(N-1),$$

which is to say $E[(\widehat{X}[n]-X[n])W^*[m]] = 0$, or equivalently $E[\widehat{X}[n]W^*[m]] = E[X[n]W^*[m]]$, for $m = n, n-1, ..., n-(N-1)$. Next, we plug in the assumed FIR form for our estimate $\widehat{X}$ to get

$$\sum_{k=0}^{N-1} h[k]\underbrace{E[W[n-k]W^*[m]]}_{R_{WW}[(n-m)-k]} = R_{XW}[n-m], \quad m = n, n-1, ..., n-(N-1).$$

Since $R_{WW}[n] = \delta[n]$ here, we get $h[n-m] = R_{XW}[n-m]$, which is equivalent to:

$$h[n] = R_{XW}[n], \quad n = 0, ..., N-1.$$

(c) In the limit as $N \nearrow \infty$, the above FIR convolution beomces unconstrained by finite order:

$$h[n] = R_{XW}[n], \quad -\infty < n < +\infty,$$

so that

$$\begin{aligned} H(\omega) &= S_{XW}(\omega) \\ &= \frac{S_{XX}(\omega)}{\sqrt{S_{YY}(\omega)}}. \end{aligned}$$

38. We have

$$Y[n] = \sum_{k_1} h[k_1]X[n-k_1],$$

so:

(a)

$$\begin{aligned} R_Y[m_1, m_2] &\triangleq E[Y[n+m_1]Y[n+m_2]Y^*[n]] \\ &= \sum_{k_1,k_2,k_3} h[k_1]h[k_2]h^*[k_3]E[Y[n+m_1-k_1]Y[n+m_2-k_2]Y^*[n]] \\ &= \sum_{k_1,k_2,k_3} h[k_1]h[k_2]h^*[k_3]R_X[m_1-k_1+k_3, m_2-k_2+k_3]. \end{aligned}$$

(b)

$$\begin{aligned} S_Y(\omega_1, \omega_2) &\triangleq \sum_{m_1,m_2} R_Y[m_1, m_2]e^{-j(\omega_1 m_1 + \omega_2 m_2)} \\ &= \sum_{m_1,m_2} \left(\sum_{k_1,k_2,k_3} h[k_1]h[k_2]h^*[k_3]R_X[m_1-k_1+k_3, m_2-k_2+k_3] \right) e^{-j(\omega_1 m_1 + \omega_2 m_2)}. \end{aligned}$$

We re-write the argument of the complex exponential $-j(\omega_1 m_1 + \omega_2 m_2)$ as follows:

$-j\omega_1 k_1 \; -j\omega_2 k_2 \; -j(m_1{-}k_1{+}k_3) \; -j(\omega_1{+}\omega_2)k_3 \; -j\left[\omega_1(m_1 - k_1 + k_3) + \omega_2(m_2 - k_2 + k_3)\right],$

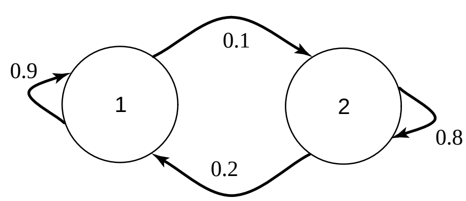
and then factor this complex exponential, to obtain

$$\begin{aligned} S_Y(\omega_1,\omega_2) &= \sum_{k_1} h[k_1]e^{-j\omega_1 k_1}\left(\sum_{k_2} h[k_2]e^{-j\omega_2 k_2}\left(\sum_{k_3} h^*[k_3]e^{+j(\omega_1+\omega_2)k_3}\left(\sum_{m_1,m_2} R_X[m_1 - k_1 + k_3, m\right.\right.\right. \\ &= \left(\sum_{k_1} h[k_1]e^{-j\omega_1 k_1}\right)\left(\sum_{k_2} h[k_2]e^{-j\omega_2 k_2}\right)\left(\sum_{k_3} h^*[k_3]e^{+j(\omega_1+\omega_2)k_3}\right) S_X(\omega_1,\omega_2) \\ &= H(\omega_1)H(\omega_2)H^*(\omega_1+\omega_2)S_X(\omega_1,\omega_2). \end{aligned}$$

39. (a) With the two states $X = 1, 2$, we have state probability vector $\mathbf{p}$ at time n (≥ 0) given as

$$\begin{aligned} \mathbf{p}[n] &= (P[X[n] = 1, P[X[n] = 2]) \\ &= (P[X[n-1] = 1]p_{11} + P[X[n-1] = 2]p_{21}, P[X[n-1] = 1]p_{12} + P[X[n-1] = 2]p_{22}) \\ &= \mathbf{p}[n-1]\mathbf{P}, \quad \text{with } \mathbf{P} \text{ the state-transition matrix,} \\ &= \mathbf{p}[n-2]\mathbf{P}^2, \\ &\;\;\vdots \\ &= \mathbf{p}[0]\mathbf{P}^n. \end{aligned}$$

(b)



(c) Let p be the probability of the event {*first* transition to state 2 occuring at time n}. Then, given $X[0] = 1$, we have

$$\begin{aligned} p &= p_{11}^{n-1}p_{12} \\ &= (0.9)^{n-1}(0.1) \\ &= (0.1)\,(0.9)^{n-1}. \end{aligned}$$

40. (a)

$$H(\omega) = \frac{1}{1 - re^{-j\omega}} \quad \text{and} \quad h[n] = r^n u[n],$$

so

$$\begin{aligned} S_{XX}(\omega) &= |H(\omega)|^2 S_{ZZ}(\omega) \\ &= \frac{1}{|1 - re^{-j\omega}|^2}\sigma_Z^2 \\ &= \frac{\sigma_Z^2}{1 + r^2 - 2r\cos\omega}. \end{aligned}$$

(b) We know $R_{XX}[m] = (h[m] * h^*[-m]) * \sigma_Z^2 \delta[m]$. Here, we have

$$\begin{aligned}
h[n] * h^*[-n] &= \sum_{k=-\infty}^{+\infty} h[k]h^*[-(n-k)] \\
&= \sum_{k=-\infty}^{+\infty} r^k u[k] r^{-(n-k)} u[k-n] \\
&= r^{-n} \sum_{k=0}^{+\infty} r^{2k} u[k-n] \\
&= \begin{cases} \frac{r^n}{1-r^2}, & n \geq 0, \\ \frac{r^{-n}}{1-r^2}, & n \leq 0, \end{cases} \\
&= \frac{r^{|n|}}{1-r^2}, \qquad \text{for all } n.
\end{aligned}$$

Thus

$$\begin{aligned}
R_{XX}[m] &= (h[m] * h^*[-m]) * \sigma_Z^2 \delta[m] \\
&= \frac{r^{|m|}}{1-r^2} * \sigma_Z^2 \delta[m] \\
&= \left(\frac{r^{|m|}}{1-r^2} * \delta[m] \right) \sigma_Z^2 \\
&= \frac{r^{|m|}}{1-r^2} \sigma_Z^2.
\end{aligned}$$

41.

42.

43. We will use the simplified notation here.

$$f_X(x_n | x_{n+1}, ..., x_{100}) = \frac{f_X(x_n, x_{n+1}, ..., x_{100})}{f_X(x_{n+1}, ..., x_{100})}. \tag{6}$$

Now, by the chain rule of probability theory, the numerator of Equation 6 is equal to

$$\begin{aligned}
& f_X(x_n, x_{n+1}) f_X(x_{n+2} | x_{n+1}, x_n) f_X(x_{n+3} | x_{n+2}, x_{n+1}, x_n) ... f_X(x_{100} | x_{99}, x_{98}, ..., x_n) \\
= \; & f_X(x_n, x_{n+1}) \prod_{k=2}^{100-n} f_X(x_{n+k} | x_{n+k-1}), \quad \text{by the Markov property.}
\end{aligned}$$

Doing a similar chain rule expansion on the denominator of Equation 6, we get

$$\begin{aligned}
& f_X(x_{n+1}) f_X(x_{n+2} | x_{n+1}) f_X(x_{n+3} | x_{n+2}, x_{n+1}) ... f_X(x_{100} | x_{99}, x_{98}, ..., x_{n+1}) \\
= \; & f_X(x_{n+1}) \prod_{k=2}^{100-n} f_X(x_{n+k} | x_{n+k-1}), \quad \text{again by the Markov property.}
\end{aligned}$$

Thus, Equation 6 becomes

$$\begin{aligned}
f_X(x_n|x_{n+1},...,x_{100}) &= \frac{f_X(x_n,x_{n+1},...,x_{100})}{f_X(x_{n+1},...,x_{100})} \\
&= \frac{f_X(x_n,x_{n+1})\prod_{k=2}^{100-n} f_X(x_{n+k}|x_{n+k-1})}{f_X(x_{n+1})\prod_{k=2}^{100-n} f_X(x_{n+k}|x_{n+k-1})} \\
&= \frac{f_X(x_n,x_{n+1})}{f_X(x_{n+1})} \\
&= f_X(x_n|x_{n+1}), \quad \text{as was to be shown.}
\end{aligned}$$

44. For the three RVs $X[n], X[n-1]$, and $X[n-2]$, we can write from basic probability theory,

$$P\left[X[n]=x[n]|X[n-2]=x[n-2]\right] = \sum_{k=-\infty}^{+\infty} P\left[X[n]=x[n], X[n-1]=x_k|X[n-2]=x[n-2]\right],$$

where the x_k are the countable set of values that may be taken on by RV $X[n-1]$. Then, by the Markov property, we can rewrite the right-hand side of this equation as

$$= \sum_{k=-\infty}^{+\infty} P\left[X[n]=x[n]|X[n-1]=x_k\right] P\left[X[n-1]=x_k|X[n-2]=x[n-2]\right].$$

By the same line of reasoning, we can write, for $n>2$,

$$\begin{aligned}
& P\left[X[n]=x[n]|X[n-2]=x[n-2], X[n-3]=x[n-3],...\right] \\
= & \sum_{k=-\infty}^{+\infty} P\left[X[n]=x[n]|X[n-1]=x_k\right] P\left[X[n-1]=x_k|X[n-2]=x[n-2], X[n-3]=x[n-3],...\right] \\
= & \sum_{k=-\infty}^{+\infty} P\left[X[n]=x[n]|X[n-1]=x_k\right] P\left[X[n-1]=x_k|X[n-2]=x[n-2]\right], \\
& \qquad \text{again by the Markov property,} \\
= & \; P\left[X[n]=x[n]|X[n-2]=x[n-2]\right], \quad \text{by first result above.}
\end{aligned}$$

45. (a) For the subsequence $X[2n]$, we can write

$$\begin{aligned}
X[2n] &= \alpha X[2n-1]+\beta W[2n] \\
&= \alpha(\alpha X[2n-2]+\beta W[2n-1])+\beta W[2n] \\
&= \alpha^2 X[2n-2]+(\alpha\beta W[2n-1+\beta W[2n]),
\end{aligned}$$

or, with $Y[n] \triangleq X[2n]$,

$$Y[n]=\alpha^2 Y[n-1]+W'[n],$$

where $W'[n] \triangleq \alpha\beta W[2n-1+\beta W[2n]$. Since $W'[n_1]$ and $W'[n_2]$ involve different values of W for all $n_1 \neq n_2$, and the fact that W is an independent random sequence, it follows that W'

is also an independent random sequence. Thus, we see that the random sequence resulting from subsampling by 2, i.e. $Y[n] = X[2n]$, is still Markov.

(b) For the variance sequence (function), we have

$$\sigma_Y^2[n] = \alpha^4 \sigma_Y^2[n-1] + \sigma_{W'}^2, \qquad n \geq 1,$$

with initial condition $\sigma_Y^2[0] = 0$. Now, $\sigma_{W'}^2 = \alpha^2\beta^2\sigma_W^2 + \beta^2\sigma_W^2 = \beta^2(1+\alpha^2)\sigma_W^2$, so the steady-state solution,

$$\begin{aligned} \sigma_Y^2[\infty] &= \frac{1}{1-\alpha^4}\beta^2(1+\alpha^2)\sigma_W^2 \\ &= \frac{\beta^2}{1-\alpha^2}\sigma_W^2. \end{aligned}$$

The homogeneous solution is given as $C\alpha^{4n}$, for some constant C, so the total solution can be written as

$$\sigma_Y^2[n] = \frac{\beta^2}{1-\alpha^2}\sigma_W^2 + C\alpha^{4n}.$$

To determine the constant, we choose $n = 1$,

$$\begin{aligned} \sigma_Y^2[1] &= \beta^2(1+\alpha^2)\sigma_W^2 \\ &= \frac{\beta^2}{1-\alpha^2}\sigma_W^2 + C\alpha^4, \end{aligned}$$

hence $C = \frac{\beta^2\sigma_W^2(1+\alpha^2-\frac{1}{1-\alpha^2})}{\alpha^4} = \frac{-\beta^2\sigma_W^2}{1-\alpha^2}$. The total solution then becomes

$$\sigma_Y^2[n] = \frac{\beta^2}{1-\alpha^2}\sigma_W^2(1-\alpha^{4n})u[n].$$

Chapter 9 solutions

1. (a) $E[X_a(t)] = E[X[n]] = \mu_X$ for $n \le t < n+1$, for all integers n. So the mean of the output analog process $\mu_{X_a}(t) = \mu_X$, a constant.

   (b) Assume times t_1 and t_2 satisfy $\boxed{\begin{array}{c} m \le t_1 < m+1, \\ \hline n \le t_2 < n+1, \end{array}}$ then

$$\begin{aligned} E[X_a(t_1)X_a(t_2)] &= R_X[m-n] \\ &= R_X[\lfloor t_1 \rfloor - \lfloor t_2 \rfloor], \end{aligned}$$

   where $\lfloor \cdot \rfloor$ is the least integer function, i.e. truncates down to the next lower integer.

2. (a)

$$\begin{aligned} \mu_X(t) &\triangleq E[X(t)] \\ &= \sum_{n=-\infty}^{+\infty} E[X[n]] \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T} \\ &= \mu_X \sum_{n=-\infty}^{+\infty} 1 \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T} \\ &= \mu_X \cdot 1 \quad \text{(since } g(t) = 1 \text{ is bandlimited with samples } g(nT) = 1) \\ &= \mu_X. \end{aligned}$$

   (b)

$$\begin{aligned} R_{XX}(t_1, t_2) &\triangleq E[X(t_1)X^*(t_2)] \\ &= \sum_{\text{all } m,n} R_{XX}[m-n] \text{sinc}\left(\frac{t_1 - mT}{T}\right) \text{sinc}\left(\frac{t_2 - nT}{T}\right), \end{aligned}$$

   where the sinc function is defined as $\text{sinc}(\tau) \triangleq \frac{\sin \pi\tau}{\pi\tau}$, so we can write the above as

$$\begin{aligned} R_{XX}(t_1, t_2) &= R_{XX}[0] \cdot \sum_n \text{sinc}\left(\frac{t_1 - nT}{T}\right) \text{sinc}\left(\frac{t_2 - nT}{T}\right) \\ &\quad + R_{XX}[1] \cdot \sum_n \text{sinc}\left(\frac{t_1 - (n+1)T}{T}\right) \text{sinc}\left(\frac{t_2 - nT}{T}\right) + \cdots \\ &= \sum_m R_{XX}[m] r_m(t_1, t_2), \end{aligned}$$

   where $r_m(t_1, t_2) \triangleq \sum_n \text{sinc}\left(\frac{t_1-(n+m)T}{T}\right) \text{sinc}\left(\frac{t_2 - nT}{T}\right)$, so then we have

$$\begin{aligned} R_{XX}(t_1, t_2) &= \sum_m R_{XX}[m] \text{sinc}\left(\frac{(t_1 - t_2) - mT}{T}\right) \quad \text{(see below)} \\ &= R_{XX}(t_1 - t_2) \quad \text{and so } X(t) \text{ is WSS.} \end{aligned}$$

   To see that $r_m(t_1, t_2) = \text{sinc}\left(\frac{(t_1-t_2)-mT}{T}\right)$, we proceed as follows: Fix m and t_1 and define the function $g(t_2) \triangleq \text{sinc}\left(\frac{(t_1-t_2)-mT}{T}\right)$. Clearly g is bandlimited with coefficients $g(nT) = \text{sinc}\left(\frac{t_1-(n+m)T}{T}\right)$, so that

$$\begin{aligned} g(t_2) &= \text{sinc}\left(\frac{(t_1-t_2)-mT}{T}\right) \\ &= \sum_n g(nT)\text{sinc}\left(\frac{t_2-nT}{T}\right) \\ &= \sum_n \text{sinc}\left(\frac{t_1-(n+m)T}{T}\right)\text{sinc}\left(\frac{t_2-nT}{T}\right). \end{aligned}$$

3. The random sequence $B[n]$ is Bernoulli and its values ± 1 occur with equal probabilities $1/2$. We have $X(t) \triangleq \sqrt{p}\sin\left(2\pi f_0 t + B[n]\frac{\pi}{2}\right)$ where $\sqrt{p}$ and f_0 are given real numbers.

(a)

$$\begin{aligned} \mu_X(t) &\triangleq E[X(t)] \\ &= E\left[\sqrt{p}\sin\left(2\pi f_0 t + B[n]\frac{\pi}{2}\right)\right] \\ &= \sqrt{p}E\left[\sin\left(2\pi f_0 t + B[n]\frac{\pi}{2}\right)\right] \\ &= \sqrt{p}\left(\frac{1}{2}\sin\left(2\pi f_0 t + \frac{\pi}{2}\right) + \frac{1}{2}\sin\left(2\pi f_0 t - \frac{\pi}{2}\right)\right) \\ &= \sqrt{p}\left(\frac{1}{2}\cos(2\pi f_0 t) + \frac{1}{2}(-\cos(2\pi f_0 t))\right) \\ &= \sqrt{p}\left(\frac{1}{2}\cos(2\pi f_0 t) - \frac{1}{2}\cos(2\pi f_0 t)\right) \\ &= 0. \end{aligned}$$

(b) For this real-valued process, $K_{XX}(t,s) = E[X(t)X(s)]$ since the means are zero. To evaluate $E[X(t)X(s)]$, we consider two cases:

(i) Case 1: $nT \le t, s < (n+1)T$, i.e. t and s are in the same half-open interval $[nT, (n+1)T)$. Then

$$\begin{aligned} E[X(t)X(s)] &= \frac{1}{2}\sqrt{p}\sin\left(2\pi f_0 t + \frac{\pi}{2}\right)\sqrt{p}\sin\left(2\pi f_0 s + \frac{\pi}{2}\right) + \frac{1}{2}\sqrt{p}\sin\left(2\pi f_0 t - \frac{\pi}{2}\right)\sqrt{p}\sin\left(2\pi f_0 s - \frac{\pi}{2}\right) \\ &= \frac{1}{2}p\cos(2\pi f_0 t)\cos(2\pi f_0 s) + \frac{1}{2}p\cos(2\pi f_0 t)\cos(2\pi f_0 s) \\ &= p\cos(2\pi f_0 t)\cos(2\pi f_0 s). \end{aligned}$$

(ii) Case 2: $nT \le t < (n+1)T, mT \le s < (m+1)T$, with $n \ne m$, i.e. t and s are in different intervals. In this case $X(t)$ and $X(s)$ are independent, so $E[X(t)X(s)] = E[X(t)]E[X(s)]$, but here the means are zero, hence $E[X(t)X(s)] = 0$.

Combining the two cases we can write

$$\begin{aligned} K_{XX}(t,s) &= E[X(t)X(s)] \\ &= \begin{cases} p\cos(2\pi f_0 t)\cos(2\pi f_0 s), & nT \le t, s < (n+1)T \text{ for some integer } n. \\ 0, & \text{else.} \end{cases} \end{aligned}$$

4.

$$Y(t) = \sum_{n=0}^{N-1} A_n X(t - nT).$$

(a)

$$\begin{aligned} E[Y(t)Y(t+\tau) &= \sum_{n,m} E[A_n A_m] E[X(t)X(t+\tau)] \\ &= \sum_{n=0}^{N-1} \sigma_A^2 R_{XX}(\tau) \qquad \text{(terms for } n \neq m \text{ are zero)} \\ &= N\sigma_A^2 R_{XX}(\tau). \end{aligned}$$

(b)

$$\begin{aligned} \Phi_Y(\omega) &\triangleq E[e^{+j\omega Y(t)}] \\ &= E\left[e^{+j\omega \sum_n A_n X(t-nT)}\right] \\ &= E\left[\prod_n \exp\left(j\omega A_n X(t-nT)\right)\right]. \end{aligned}$$

Now the A_n are jointly independent and also independent of the random process $X(t)$. Also $R_{XX}(nT) = 0$, all of the $X(t - nT)$ for fixed t, are also jointly independent, since Gaussian. Therefore, all the terms in the above product are jointly independent, thus

$$\Phi_Y(\omega) = \prod_n E[\exp\left(j\omega A_n X(t-nT)\right)].$$

Next, use conditional expectation to write $E[\exp\left(j\omega A_n X(t-nT)\right)] = E[E[\exp\left(j\omega A_n X(t-nT)\right) | A_n]]$, where $E[\exp\left(j\omega A_n X(t-nT)\right) | A_n = a_n] = \exp(-\frac{1}{2}R_{XX}(0)\omega^2 a_n^2)$. Then we have

$$\begin{aligned} E[\exp\left(j\omega A_n X(t-nT)\right)] &= E[\exp(-\frac{1}{2}R_{XX}(0)\omega^2 A_n^2)] \\ &= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\sigma_X^2\omega^2 a^2} \frac{1}{\sqrt{2\pi\sigma_A^2}} e^{-\frac{1}{2}\frac{a^2}{\sigma_A^2}}\, da \qquad \text{(with } \sigma_X^2 = R_{XX}(0)\text{)} \\ &= \frac{1}{\sqrt{2\pi\sigma_A^2}} \int_{-\infty}^{+\infty} e^{-\frac{a^2}{2}\left(\sigma_X^2\omega^2 + \frac{1}{\sigma_A^2}\right)} da \\ &= \left(\sigma_X^2\sigma_A^2\omega^2 + 1\right)^{-\frac{1}{2}}. \end{aligned}$$

Hence

$$\Phi_Y(\omega) = \left(\sigma_X^2\sigma_A^2\omega^2 + 1\right)^{-\frac{N}{2}}.$$

(c)

$$\begin{aligned} \Phi_Y\left(\frac{\omega}{\sqrt{N}}\right) &= \frac{1}{\left(\sigma_X^2\sigma_A^2\frac{\omega^2}{N} + 1\right)^{\frac{N}{2}}} \\ &\simeq \left(e^{-\frac{\sigma_X^2\sigma_A^2\omega^2}{N}}\right)^{N/2} \\ &= e^{-\frac{1}{2}\sigma_X^2\sigma_A^2\omega^2}. \end{aligned}$$

Hence the asymptotic CDF of $\frac{1}{\sqrt{N}}Y(t)$, as $N \longrightarrow \infty$, is Normal (Gaussian): $N(0, \sigma_X^2\sigma_A^2)$.

(d) The number of taps N is Poisson now, so we must re-do (a) and (b). First

$$E[Y(t+\tau)Y(t)|N] = N\sigma_A^2 R_{XX}(\tau),$$

which implies

$$\begin{aligned} E[Y(t+\tau)Y(t)] &= E[N\sigma_A^2 R_{XX}(\tau)] \\ &= E[N]\sigma_A^2 R_{XX}(\tau) \\ &= \lambda\sigma_A^2 R_{XX}(\tau). \end{aligned}$$

Then

$$\begin{aligned} \Phi_Y(\omega) &= E[e^{+j\omega Y(t)}] \\ &= E[E[e^{+j\omega Y(t)}|N]] \\ &= E\left[\left(\frac{1}{\sigma_X^2\sigma_A^2\omega^2+1}\right)^{N/2}\right] \\ &= \sum_{n=0}^{\infty} d^{n/2}e^{-\lambda}\frac{\lambda^n}{n!}, \qquad \text{where} \quad d \triangleq \frac{1}{\sigma_X^2\sigma_A^2\omega^2+1}, \\ &= \left(\sum_{n=0}^{\infty}\frac{(\lambda\sqrt{d})^n}{n!}\right)e^{-\lambda} \\ &= \exp(\lambda\sqrt{d}-\lambda) \\ &= \exp(\lambda(\sqrt{d}-1)). \end{aligned}$$

5. By definition of Poisson process with parameter $\lambda(>0)$, we have

$$\begin{aligned} P_N(n;t) &= P[N(t)=n] \\ &= \frac{\lambda t}{n!}e^{-\lambda t}u[n], \end{aligned}$$

where $u[n]$ is the unit-step function.

(a) Let $t_2 \geq t_1$, then be independent increments property, $N(t_2)-N(t_1)$ and $N(t_1)$ are independent RVs. Also the increment is Poisson distributed with the same parameter λ. Hence

$$\begin{aligned} P_N(n_1,n_2;t_1,t_2) &= P[N(t_1)=n_1, N(t_2)=n_2] \\ &= P[N(t_1)=n_1]P[N(t_2)-N(t_1)=n_2-n_1] \\ &= \frac{\lambda t_1}{n_1!}e^{-\lambda t_1}u[n_1]\frac{\lambda(t_2-t_1)}{(n_2-n_1)!}e^{-\lambda(t_2-t_1)}u[n_2-n_1] \\ &= \frac{\lambda t_1}{n_1!}e^{-\lambda t_1}\frac{\lambda(t_2-t_1)}{(n_2-n_1)!}e^{-\lambda(t_2-t_1)}u[n_1]u[n_2-n_1] \\ &= \frac{\lambda^2 t_1(t_2-t_1)}{n_1!(n_2-n_1)!}e^{-\lambda t_2}u[n_1]u[n_2-n_1]. \end{aligned}$$

(b) Since the t_i's are increasing, $N(t)$ has independent increments, which can be recursively applied to conclude

$$\begin{aligned} P_N(n_1, n_2, ..., n_K; t_1, t_2, ..., t_K) &= P[N(t_1) = n_1, N(t_2) = n_2, ..., N(t_K) = n_K] \\ &= P[N(t_1) = n_1]P[N(t_2) - N(t_1) = n_2 - n_1] \cdots \\ &\qquad \cdots P[N(t_K) - N(t_{K-1}) = n_K - n_{K-1}] \\ &= \frac{\lambda^K t_1(t_2 - t_1) \cdots (t_K - t_{K-1})}{n_1!(n_2 - n_1)! \cdots (n_K - n_{K-1})!} e^{-\lambda t_K} u[n_1]u[n_2 - n_1] \cdots u[n_K - n_{K-1}]. \end{aligned}$$

6. (a) Use property (3) for $t_1 = 0$ and $t_2 = t$. Then by the property (1), $N(0) = 0$. So, (3) becomes:

$$P[N(t) = n] = \frac{\left(\int_0^t \lambda(s)ds\right)^n}{n!} e^{-\left(\int_0^t \lambda(s)ds\right)^n} \quad \text{for } n \geq 0.$$

Then since $N(t)$ is Poisson distributed, we recognize the mean as

$$\mu_N(t) = \int_0^t \lambda(s)ds, \qquad t \geq 0.$$

(b) Take $t_2 \geq t_1 \geq 0$, and write $E[N(t_1)N(t_2)] = E[N(t_1)[N(t_1) + (N(t_2) - N(t_1))]]$. Then using the linearity of the expectation operator E and the independent increments property (2), we get

$$\begin{aligned} R_N(t_1, t_2) &\triangleq E[N(t_1)N(t_2)] \\ &= E[N^2(t_1)] + E[N(t_1)]E[N(t_2) - N(t_1)]. \end{aligned}$$

We then recognize the first term on the rhs as the second moment of the Poisson, therefore

$$E[N^2(t_1)] = \int_0^t \lambda(s)ds + \left(\int_0^t \lambda(s)ds\right)^2.$$

Now $E[N(t_2) - N(t_1)] = \int_{t_1}^{t_2} \lambda(s)ds$, and from part (a) $E[N(t_1)] = \int_0^t \lambda(s)ds$, so, putting these together, we get

$$\begin{aligned} R_N(t_1, t_2) &= E[N^2(t_1)] + E[N(t_1)]E[N(t_2) - N(t_1)], \qquad t_2 \geq t_1 \geq 0, \\ &= \int_0^{t_1} \lambda(s)ds + \left(\int_0^{t_1} \lambda(s)ds\right)^2 + \left(\int_0^{t_1} \lambda(s)ds\right)\left(\int_{t_1}^{t_2} \lambda(s)ds\right) \\ &= \left(\int_0^{t_1} \lambda(s)ds\right)\left(1 + \int_0^{t_2} \lambda(s)ds\right), \qquad t_2 \geq t_1 \geq 0. \end{aligned}$$

For the general case, from the symmetry of the correlation function $R_N(t_1, t_2) \triangleq E[N(t_1)N(t_2)]$, we can write

$$R_N(t_1, t_2) = \left(\int_0^{\min(t_1, t_2)} \lambda(s)ds\right)\left(1 + \int_0^{\max(t_1, t_2)} \lambda(s)ds\right), \qquad t_1, t_2 \geq 0.$$

(c) We have to show properties (1), (2), and (3):

(1) $N_u(0) \triangleq N(t(0)) = N(0) = 0$. ✓

(2) Let $\tau_1 \leq \tau_2 \leq \tau_3 \leq \cdots \leq \tau_k$. Then since $t(\tau)$ is monotone increasing (since it is the integral of a positive $\lambda(s)$), we have $t_1 \leq t_2 \leq t_3 \leq \cdots \leq t_k$ where $t_i \triangleq t(\tau_i)$. Thus $N_u(\tau_i) \triangleq N(t(\tau_i)) = N(t_i)$. So, by definition, $N(t_1), N(t_2) - N(t_1), ...,$ $N(t_k) - N(t_{k-1})$ are jointly independent. But $N_u(\tau_i) - N_u(\tau_{i-1}) = N(t_i) - N(t_{i-1})$, so the $N_u(\tau)$ process also has independent increments.

(3) Since $N_u(\tau_2) - N_u(\tau_1) = N(t_2) - N(t_1)$ with mean value

$$\begin{aligned}\int_{t_1}^{t_2} \lambda(s)ds &= \int_0^{t_2} \lambda(s)ds - \int_0^{t_1} \lambda(s)ds \\ &= \tau_2 - \tau_1\end{aligned}$$

which means that $N(\tau)$ has λ parameter equal to 1.

7. The Poisson PMF is given as

$$P_N(n;t) = \frac{\left(\int_0^t \lambda(\nu)d\nu\right)^n}{n!} e^{-\left(\int_0^t \lambda(\nu)d\nu\right)^n} u[n].$$

(a)

$$\begin{aligned}\mu_N(t) &= E[N(t)] \\ &= \int_0^t \lambda(\nu)d\nu, \qquad t \geq 0, \\ &= \int_0^t (1+2\nu)d\nu \\ &= \nu + \nu^2\big|_0^t \\ &= t + t^2, \qquad t \geq 0.\end{aligned}$$

(b)Let $t_2 \geq t_1 \geq 0$, then

$$N(t_1)N(t_2) = N(t_1)[N(t_1) + (N(t_2) - N(t_1))],$$

so

$$R_N(t_1,t_2) = E[N^2(t_1)] + \mu_N(t_1)\left(\mu_N(t_2) - \mu_N(t_1)\right).$$

Now

$$\begin{aligned}E[N^2(t)] &= \int_0^t \lambda(\nu)d\nu + \left(\int_0^t \lambda(\nu)d\nu\right)^2 \\ &= \left(t+t^2\right) + \left(t+t^2\right)^2.\end{aligned}$$

So for $t_2 \geq t_1$,

$$\begin{aligned}R_N(t_1,t_2) &= \left(t_1+t_1^2\right) + \left(t_1+t_1^2\right)^2 + \left(t_1+t_1^2\right)\left(\left(t_2+t_2^2\right) - \left(t_1+t_1^2\right)\right) \\ &= \left(t_1+t_1^2\right) + \left(t_1+t_1^2\right)\left(t_2+t_2^2\right).\end{aligned}$$

In general, we have

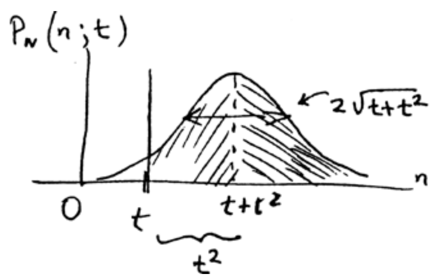
$$R_N(t_1,t_2) = \left[\min(t_1,t_2) + (\min(t_1,t_2))^2\right]\left[1 + \max(t_1,t_2) + (\max(t_1,t_2))^2\right].$$

(c)

$$\begin{aligned} P[N(t) \geq t] &= \sum_{n\geq \lceil t \rceil} P_N(n;t) \\ &= \sum_{n\geq \lceil t \rceil} \frac{(t+t^2)^n}{n!} e^{-(t+t^2)}, \qquad t \geq 0. \end{aligned}$$

(d) Use the CLT with $\mu_N(t) = t + t^2$ and $\sigma_N^2 = t + t^2$ to yield

$$\begin{aligned} P[N(t) &\geq t] \approx \frac{1}{2} + \operatorname{erf}\left(\frac{t^2}{\sqrt{t+t^2}}\right) \\ &\approx \frac{1}{2} + \operatorname{erf}(t). \end{aligned}$$



We remember

$$\begin{aligned} \operatorname{erf}(x) &= \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}v^2} dv \\ &= P[X_{SN} \leq x] \quad \text{for} \quad X_{SN} : N(0,1). \end{aligned}$$

8. (a) We know

$$f_T(t;n) = \underbrace{f_T(t) * f_T(t) * \cdots f_T(t)}_{n \text{ times}},$$

with $f_T(t) = \lambda e^{-\lambda t} u(t)$. Therefore

$$\begin{aligned} f_T(t;2) &= f_T(t) * f_T(t) \\ &= \left(\int_0^t \lambda e^{-\lambda \tau} \lambda e^{-\lambda(t-\tau)} d\tau\right) u(t) \\ &= \lambda^2 e^{-\lambda t} \left(\int_0^t d\tau\right) u(t) \\ &= \lambda^2 t e^{-\lambda t} u(t). \end{aligned}$$

For $n = 3$, we have

$$\begin{aligned} f_T(t;3) &= f_T(t;2) * f_T(t) \\ &= \lambda^2 t e^{-\lambda t} u(t) * \lambda e^{-\lambda t} u(t) \\ &= \lambda^3 \left(\int_0^t \tau e^{-\lambda \tau} e^{-\lambda (t-\tau)} d\tau \right) u(t) \\ &= \lambda^3 e^{-\lambda t} \left(\int_0^t \tau d\tau \right) u(t) \\ &= \lambda^3 \frac{t^2}{2} e^{-\lambda t} u(t). \end{aligned}$$

For general n, we compute $f_T(t;n) = f_T(t;n-1) * f_T(t)$, which can similarly be seen to be satisfied by

$$f_T(t;n) = \lambda^n \frac{t^{n-1}}{n!} e^{-\lambda t} u(t).$$

(b)

$$\begin{aligned} F_{\tau'[i]}(\tau'|\tau[i] &\geq t) = P[\tau[i] \leq \tau' + t | \tau[i] \geq t] \\ &= \frac{\int_t^{t+\tau'} \lambda e^{-\lambda \tau} d\tau}{\int_t^{\infty} \lambda e^{-\lambda \tau} d\tau} \\ &= \frac{-e^{-\lambda \tau}|_t^{t+\tau'}}{-e^{-\lambda \tau}|_t^{\infty}} \\ &= 1 - e^{-\lambda \tau'}, \end{aligned}$$

thus, by differentiation, $f_{\tau'[i]}(\tau') = |\tau[i] \geq t) = \lambda e^{-\lambda \tau'} u(\tau')$, the same density as τ.

(c)

$$f_{\tau'[i]}(\tau') = \int_{-\infty}^{+\infty} f_{\tau'[i]}(\tau'|T = t) f_T(t) dt,$$

but from part (b), the conditional pdf $f_{\tau'[i]}(\tau') = |\tau[i] \geq t)$ is independent of the variable t, so

$$\begin{aligned} f_{\tau'[i]}(\tau') &= f_{\tau'[i]}(\tau'|T = t) \int_{-\infty}^{+\infty} f_T(t) dt \\ &= f_{\tau'[i]}(\tau'|T = t) \\ &= f_{\tau[i]}(\tau'), \qquad \text{using the result of (b).} \end{aligned}$$

9. (a) Use the general conditional expectation property $E[X] = E[E[X|Y]]$ to conclude

$$\begin{aligned} \mu_N(t) &= E[N(t)] \\ &= E[E[N(t)|S(t)]] \\ &= E[\int_0^t S(\tau) d\tau] \\ &= \int_0^t \mu_S(\tau) d\tau \\ &= \mu_0 t. \end{aligned}$$

(b)

$$
\begin{aligned}
\sigma_N^2(t) &= E[N^2(t)] - (E[N(t)])^2 \\
&= E[E[N^2(t)|S(t)]] - (\mu_0 t)^2 \\
&= E\left[\int_0^t S(\tau)d\tau + \left(\int_0^t S(\tau)d\tau\right)^2\right] - (\mu_0 t)^2 \\
&= \mu_0 t + \int_0^t \int_0^t K_S(\tau_1, \tau_2) d\tau_1 d\tau_2.
\end{aligned}
$$

10.

$$
\begin{aligned}
P_K(k;t) &= \sum_{n=k}^{\infty} P[K(t) = k, N(t) = n] \\
&= \sum_{n=k}^{\infty} P[K(t) = k|N(t) = n] P[N(t) = n] \\
&= \sum_{n=k}^{\infty} \binom{n}{k} p^k q^{n-k} \cdot \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad k \geq 0, t \geq 0, \\
&= \sum_{n'=0}^{\infty} \frac{p^k q^{n'}}{n'!k!} (\lambda t)^{n'+k} e^{-\lambda t} \quad \text{with } n' \triangleq n - k, \\
&= \left(\sum_{n'=0}^{\infty} \frac{(q\lambda t)^{n'}}{n'1}\right) \frac{p^k (\lambda t)^k}{k!} e^{-\lambda t} \\
&= e^{q\lambda t} \frac{(p\lambda t)^k}{k!} e^{-\lambda t} \\
&= \frac{(p\lambda t)^k}{k!} e^{-p\lambda t}, \quad k \geq 0, t \geq 0,
\end{aligned}
$$

a Poisson RV with parameter $p\lambda t$.

11. (a) Use conditional expectation and write

$$
\begin{aligned}
E[N(x)] &= E[E\{N(x)|S(x)]] \\
&= E[\lambda(x)] \\
&= E[S(x) + \lambda_0]T \\
&= (E[S(x)] + \lambda_0) T \\
&= \lambda_0 T.
\end{aligned}
$$

Then for the variance, we have

$$
\begin{aligned}
E\left[(N(x) - E[N(x)])^2\right] &= E\left[E[(N(x) - E[N(x)])^2]|S(x)\right] \\
&= \sigma_S^2 T^2 + \lambda_0 T.
\end{aligned}
$$

(b)

$$\begin{aligned}
R_N(x_1,x_2) &\triangleq E[N(x_1)N(x_2)] \\
&= E[E[N(x_1)N(x_2)|S(x_1),S(x_2)]] \\
&= E[E[N(x_1)|S(x_1)]E[N(x_2)|S(x_2)] \quad \text{by conditional independence of } N \text{ given } S, \\
&= E[(S(x_1)] + \lambda_0)\,T\,(S(x_2)] + \lambda_0)\,T] \\
&= E[S(x_1)S(x_2)]T^2 + (\lambda_0 T)^2 \qquad \text{since } S \text{ is zero mean,} \\
&= \left(\sigma_S^2 \exp(-\alpha|x_1 - x_2|) + \lambda_0^2\right) T^2.
\end{aligned}$$

12. (a) In each 'state,' the time to the next transition is governed by the *interarrival time sequence* $\tau[n]$ for the Poisson process. These are i.i.d. exponentially distributed RVs. Hence, future behavior of $X(t)$ only depends on this current value, the state.

(b) For δ small and positive, we write

$$P[X(t) = +1] = P[X(t-\delta) = +1](1 - \lambda\delta) + P[X(t-\delta) = -1]\lambda\delta,$$

so in steady state, letting $P_1 \triangleq P_X(1,\infty)$ and $P_{-1} \triangleq P_X(-1,\infty)$, where $P_{-1} = 1 - P_1$, we have $P_1 = P_1(1-\lambda\delta) + P_{-1}\lambda\delta$ or $P_1\lambda\delta = P_{-1}\lambda\delta$. Hence $P_1 = 1/2$.

(c) Rewriting the first equation in (b) using PMF notation, we have

$$P_X(1;t) - P_X(1;t-\delta) = -\lambda P_X(1;t-\delta)\delta + \lambda P_X(-1;t-\delta)\delta.$$

Dividing this equation by δ, and taking limits as $\delta \to 0$, we get

$$\frac{dP_X(1;t)}{dt} = -\lambda P_X(1;t) + \lambda P_X(-1;t),$$

and similarly we obtain

$$\frac{dP_X(-1;t)}{dt} = -\lambda P_X(-1;t) + \lambda P_X(1;t).$$

13. (a)

$$\begin{aligned}
P_N(n;t) &= P[T[n] \le t, T[n+1\} > t], \qquad n \ge 1, \\
&= P[\{T[n] \le t\} - \{T[n+1] \le t\}], \text{ where minus sign indicates } \textit{event subtraction}, \\
&= P[[T[n] \le t] - P[T[n+1] \le t] \\
&= F_T(t;n) - F_T(t;n+1).
\end{aligned}$$

(b)

$$P_N(0;t) = \begin{cases} 1-t, & 0 \le t \le 1, \\ 0, & t > 1. \end{cases}$$

$$P_N(1;t) = \begin{cases} t - \frac{t^2}{2}, & 0 \le t \le 1, \\ 2 - 2t + \frac{t^2}{2}, & 1 < t \le 2 \\ 0 & t > 2. \end{cases}$$

These results come from

$$F_T(t;1) = \begin{cases} t, & 0 \le t \le 1, \\ 1, & t > 1, \end{cases}$$

and

$$F_T(t;2) = \begin{cases} \frac{t^2}{2}, & 0 \le t \le 1, \\ -1 + 2t - \frac{t^2}{2}, & 1 < t \le 2, \\ 1, & t > 2, \end{cases}$$

which, in turn, come from running integration, i.e. $\int_{-\infty}^{t}$, on the pdf's

$$\begin{aligned} f_T(t;1) &= f_\tau(t) \\ &= u(t) - u(t-1), \end{aligned}$$

and

$$f_T(t;2) = \begin{cases} t, & 0 \le t \le 1, \\ 2 - t, & 1 < t \le 2. \end{cases}$$

For $P_N(2;t)$, we get

$$P_N(2;t) = \begin{cases} \frac{t}{2} - \frac{t^3}{6}, & 0 \le t \le 1, \\ -\frac{3}{2} + \frac{7}{2}t - 2t^2 + \frac{t^3}{3}, & 1 < t \le 2, \\ \frac{9}{2} - \frac{9}{2}t + \frac{3}{2}t^2 - \frac{t^3}{3}, & 2 < t \le 3, \\ 0, & t > 3, \end{cases}$$

which is supported by the pdf

$$f_T(t;3) = \begin{cases} \frac{t^2}{2}, & 0 \le t \le 1, \\ -\frac{3}{2} + 3t - t^2, & 1 \le t \le 2, \\ \frac{9}{2} - 3t + \frac{t^2}{2}, & 2 < t \le 3, \\ 0, & t > 3, \end{cases}$$

and CDF

$$F_T(t;3) = \begin{cases} \frac{1}{6}t^3, & 0 \le t \le 1, \\ \frac{1}{2} - \frac{3}{2}t + \frac{3}{2}t^2 - \frac{t^3}{3}, & 1 \le t \le 2, \\ -\frac{7}{2} + \frac{9}{2}t - \frac{3}{2}t^2 + \frac{t^3}{6}, & 2 < t \le 3, \\ 1, & t > 3. \end{cases}$$

(c) We have

$$T[n] = \sum_{k=1}^{n} \tau[k], \quad \text{and CF} \quad \Phi_\tau(\omega) = \left(\frac{\sin \omega/2}{\omega/2}\right) e^{+j\omega/2} \;,$$

so

$$\begin{aligned} \Phi_{T[n]}(\omega) &= (\Phi_\tau(\omega))^n \\ &= \left(\frac{\sin \omega/2}{\omega/2}\right)^n e^{+j\omega n/2}. \end{aligned}$$

Note: In the 1st printing, part (c) asks for the CF of the renewal process $N(t)$, but the CF of the interarrival time sequence $T[n]$ is all that is needed in part (d) of this problem, in order to obtain an approximate expression for the PMF $P_N(n;t)$ via the result of part (a).

(d) Now, since $\left(\frac{\sin \omega/2}{\omega/2}\right) \approx 1 - \frac{1}{6}\left(\frac{\omega}{2}\right)^2$ for small ω, which in turn is $\approx \exp(-\frac{1}{6}\left(\frac{\omega}{2}\right)^2)$, via $e^{-x} \approx 1 - x$, we get

$$\begin{aligned} \Phi_{T[n]}(\omega) &= \left(\frac{\sin \omega/2}{\omega/2}\right)^n e^{+j\omega n/2} \\ &\approx \exp(-\frac{n}{12}\omega^2 + j\omega n/2), \end{aligned}$$

which is Normal with mean $n/2$ and variance $n/12$, i.e. $T[n] : N(n/2, n/12)$. Thus

$$F_T(t; n) \approx \begin{cases} \frac{1}{2} - \operatorname{erf}\left(\frac{n/2-t}{\sqrt{n/12}}\right), & t \leq n/2, \\ \frac{1}{2} + \operatorname{erf}\left(\frac{t-n/2}{\sqrt{n/12}}\right), & t > n/2, \end{cases}$$

where the approximation is good near the mean, i.e. within a few std's, and for large n.

14. The standard Wiener process $W(t)$ is distributed as $N(0, t)$ and defined on $[0, \infty)$. Using the independent increments property as in problem 9.5, we have for any $t_2 > t_1$,

$$\begin{aligned} f_W(a_1, a_2; t_1, t_2) &= f_W(a_1; t_1) f_W(a_2 - a_1; t_2 - t_1) \\ &= \frac{1}{\sqrt{2\pi t_1}} \exp\left(-\frac{1}{2}\frac{a_1^2}{t_1}\right) \frac{1}{\sqrt{2\pi (t_2 - t_1)}} \exp\left(-\frac{1}{2}\frac{(a_2 - a_1)^2}{t_2 - t_1}\right). \end{aligned}$$

15. Since W_1 and W_2 are independent Wiener processes, there difference is Gaussian distributed. Since the mean of $W_2(t)$ is zero, $-W_2(t)$ still has the correlation function $\alpha_2 \min(t_1, t_2)$. Thus

(a)

$$R_X(t_1, t_2) = (\alpha_1 + \alpha_2) \min(t_1, t_2)$$

and

(b)

$$X(t) : N(0, (\alpha_1 + \alpha_2)t).$$

16. (a) $Y(t) = X'(t)$, so $\mu_Y(t) = \frac{d}{dt}\mu_X(t) = \frac{d}{dt}(4) = 0$.

(b) Since $\mu_X(t) = 4$,

$$\begin{aligned} R_{YY}(t_1, t_2) &= \frac{\partial^2}{\partial t_1 \partial_{t_2}}[5 \min{}^2(t_1, t_2) + 4] \\ &= \frac{\partial}{\partial t_1}\left[\frac{\partial}{\partial_{t_2}}\left(5 \min{}^2(t_1, t_2)\right)\right] \\ &= \frac{\partial}{\partial t_1}[10 \min(t_1, t_2) u(t_1 - t_2)] \\ &= 10 \min(t_1, t_2) \delta(t_1 - t_2), \end{aligned}$$

where we have written 0 for $u(t_2 - t_0)u(t_1 - t_2)$, since it is only non-zero at the one point $t_1 = t_2$, and there only takes on the finite value 1. So

$$\begin{aligned} K_{YY}(t_1, t_2) &= R_{YY}(t_1, t_2) \\ &= 10 \min(t_1, t_2) \delta(t_1 - t_2) \\ &= 10 t_1 \delta(t_1 - t_2) = 10 t_2 \delta(t_1 - t_2). \end{aligned}$$

(c) (i) $\mu_Y(t) = 0$.

(ii) $R_{YY}(t + \tau, t) = (t + \tau)\delta(\tau) = t\delta(\tau)$, not WSS.

(d) One way to get this covariance is to multiply together two jointly independent Wiener processes $W_1(t)$ and $W_2(t)$, and then scale the result by $\sqrt{5}$. Call the resulting process

$$X_c(t) \triangleq \sqrt{5} W_1(t) W_2(t).$$

Since the W_i have covariance function $\min(t_1, t_2)$, we get

$$
\begin{aligned}
K_{X_cX_c}(t_1,t_2) &= E[\sqrt{5}W_1(t_1)W_2(t_1)\sqrt{5}W_1(t_2)W_2(t_2)] \\
&= 5E[W_1(t_1)W_1(t_2)W_2(t_1)W_2(t_2)] \\
&= 5E[W_1(t_1)W_1(t_2)]E[W_2(t_1)W_2(t_2)] \\
&= 5\min(t_1,t_2)\min(t_1,t_2) \\
&= 5\min{}^2(t_1,t_2).
\end{aligned}
$$

17.

18. $X(t)$ is a Markov random process on $[0,\infty)$ with $f_X(x;0) = \delta(x-1)$ and

$$
f_X(x_2|x_1;t_2,t_1) = \frac{1}{\sqrt{2\pi\,(t_2-t_1)}}\exp\left(-\frac{(x_2-x_1)^2}{2\,(t_2-t_1)}\right), \qquad t_2 > t_1.
$$

(a) We know

$$
\begin{aligned}
f_X(x_2;t_2) &= \int_{-\infty}^{+\infty} f_X(x_2|x_1;t_2,0)f_X(x_1;0)dx_1 \\
&= \frac{1}{\sqrt{2\pi t_2}}\exp\left(-\frac{(x_2-1)^2}{2t_2}\right),
\end{aligned}
$$

so $f_X(x;t) = \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(x-1)^2}{2t}\right)$ for all $t > 0$.

(b) Here we are given $X(0) : N(0,1)$, thus, following the method in (a), we have

$$
\begin{aligned}
f_X(x_2;t_2) &= \int_{-\infty}^{+\infty} f_X(x_2|x_1;t_2,0)f_X(x_1;0)dx_1 \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t_2}}\exp\left(-\frac{(x_2-x_1)^2}{2t_2}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x_1^2}{2}\right)dx_1.
\end{aligned}
$$

By completing the square (see math. Appendix A), we have

$$
\begin{aligned}
f_X(x_2;t_2) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^2\,t_2}}\exp\left(-\frac{x_2^2}{2(t_2+1)} - \frac{(t_2+1)\left(x_1-\frac{x_2}{t_2+1}\right)^2}{2}\right)dx_1 \\
&= \frac{1}{\sqrt{2\pi(t_2+1)}}\exp\left(-\frac{x_2^2}{2(t_2+1)}\right),
\end{aligned}
$$

which is the same as $f_X(x;t) = \frac{1}{\sqrt{2\pi(t+1)}}\exp\left(-\frac{x^2}{2(t+1)}\right)$, i.e. $X(t) : N(0,t+1)$, for all $t > 0$.

19. (a)

$$
\begin{aligned}
P[\textit{remain}\text{ in state 2 until time } t|X(0) &= 2] = P[\text{no transition to 3}]P[\text{no transition to 1}] \\
&= e^{-\lambda_2 t}e^{-\mu_2 t} \\
&= \exp-(\lambda_2+\mu_2)t.
\end{aligned}
$$

(b)

$$\begin{bmatrix} P_1(t+\Delta t) \\ P_2(t+\Delta t) \\ P_3(t+\Delta t) \end{bmatrix} \approx \begin{bmatrix} 1-\lambda_1\Delta t & \mu_2\Delta t & 0 \\ \lambda_1\Delta t & 1-(\lambda_2+\mu_2)\Delta t & \mu_3\Delta t \\ 0 & \lambda_2\Delta t & 1-\mu_3\Delta t \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{bmatrix}.$$

Then taking the limit as goes to zero, we get the vector equations

$$\frac{d\mathbf{P}}{dt} = \begin{bmatrix} -\lambda_1 & \mu_2 & 0 \\ \lambda_1 & -(\lambda_2+\mu_2) & \mu_3 \\ 0 & \lambda_2 & -\mu_3 \end{bmatrix} \mathbf{P}(t),$$

where $\mathbf{P}(t) \triangleq (P_1(t), P_2(t), P_3(t))^T$.

(c) In the steady state $d\mathbf{P}/dt = 0$ and from the above vector equation we obtain

$$\begin{aligned} -\lambda_1 P_1 + \mu_2 P_2 &= 0 \quad \text{and} \\ \lambda_2 P_2 - \mu_3 P_3 &= 0. \end{aligned}$$

Solving these two, we obtain

$$P_2 = \frac{\lambda_1}{\mu_2} P_1 \qquad \text{and} \qquad P_3 = \frac{\lambda_2}{\mu_3} P_2.$$

Now, together with $P_1 + P_2 + P_3 = 1$, which always holds, we get

$$\begin{aligned} P_1 + P_2 + P_3 &= P_1 + \frac{\lambda_1}{\mu_2} P_1 + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2} P_1 \\ &= P_1 \left(1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}\right) \\ &= 1, \end{aligned}$$

so,

$$\begin{aligned} P_1 &= \frac{1}{1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}, \\ P_2 &= \frac{\frac{\lambda_1}{\mu_2}}{1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}, \\ P_3 &= \frac{\frac{\lambda_1}{\mu_2}\frac{\lambda_2}{\mu_3}}{1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}. \end{aligned}$$

20.

21. (a)

$$f_X(x(t)) = \frac{1}{\sqrt{2\pi\alpha t}} \exp\left(-\frac{x(t)^2}{2\alpha t}\right), \qquad -\infty < x(t) < +\infty.$$

(b)

$$f_X(x(t)|x(s)) = \frac{1}{\sqrt{2\pi\alpha(t-s)}} \exp\left(-\frac{(x(t)-x(s))^2}{2\alpha(t-s)}\right), \qquad t > s.$$

(c)

$$\frac{1}{\sqrt{2\pi 2\alpha\Delta}} \exp\left(-\frac{(x_3-x_1)^2}{2\alpha 2\Delta}\right) = \frac{1}{2\pi\alpha\Delta}\int_{-\infty}^{+\infty} \exp\left(-\frac{(x_3-x_2)^2+(x_2-x_1)^2}{2\alpha\Delta}\right) dx_2.$$

We then complete the square as

$$(x_3-x_2)^2+(x_2-x_1)^2 = 2\left[\left(x_2-\frac{x_1+x_3}{2}\right)^2+\left(\frac{x_1-x_3}{2}\right)^2\right],$$

so

$$\begin{aligned}\frac{1}{\sqrt{2\pi 2\alpha\Delta}} \exp\left(-\frac{(x_3-x_1)^2}{2\alpha 2\Delta}\right) &= \frac{\sqrt{2\pi\alpha\Delta/2}}{2\pi\alpha\Delta}\exp\left(-\frac{2(x_1-x_3)^2/4}{2\alpha\Delta}\right)\\ &= \frac{1}{\sqrt{2\pi 2\alpha\Delta}}\exp\left(-\frac{(x_1-x_3)^2}{4\alpha\Delta}\right).\end{aligned}$$

22. The covariance function is just a function of the difference of the two times t and s. Also the mean $\mu_{X'} = 0$ is constant. This satisfies the definition of wide sense stationarity. For (strict sense) stationarity, one would need to know that all higher order moment functions were also shift invariant. So, we cannot conclude stationarity here.

23. (a)

$$\begin{aligned}\mu_X(t) &= \mu_A \cos 2\pi ft + \mu_B \sin 2\pi ft\\ &= 0\cdot\cos 2\pi ft + 0\cdot \sin 2\pi ft\\ &= 0.\end{aligned}$$

$$K_{XX}(t,s) = E[X(t)X(s)]$$

$$\begin{aligned}&= E[(A\cos 2\pi ft + B\sin 2\pi ft)(A\cos 2\pi fs + B\sin 2\pi fts)]\\ &= E[A^2]\cos 2\pi ft\cos 2\pi fs + E[B^2]\sin 2\pi ft \sin 2\pi fts\\ &= \sigma^2(\cos 2\pi ft\cos 2\pi fs + \sin 2\pi ft\sin 2\pi fts)\\ &= \sigma^2\cos 2\pi f(t-s).\end{aligned}$$

(b) It can be shown that $E[X^3(t)]$ is not constant, as follows.

$$\begin{aligned}E[X^3(t)] &= E[(A\cos 2\pi ft + B\sin 2\pi ft)^3]\\ &= E[A^3\cos^3 2\pi ft + 3A^2B\cos^2 2\pi ft\sin 2\pi ft\\ &\quad +3AB^2\cos 2\pi ft\sin^2 2\pi ft + B^3\sin^3 2\pi ft]\\ &= \mu(\cos^3 2\pi ft + \sin^3 2\pi ft)\\ &\neq \text{constant}.\end{aligned}$$

The second to last line is because $E[A^2B] = E[A^2]E[B] = 0$ and $E[AB^2] = E[A]E[B^2] = 0$, since $E[A] = E[B] = 0$. We also write $\mu = E[A^3] = E[B^3]$. The last line is because $\mu \neq 0$ and $\cos^3 2\pi ft + \sin^3 2\pi ft$ is not a constant with respect to time.

24. We are asked to prove that

$$K_{YY}(t_1,t_2) = L_{t_1}L_{t_2}\{K_{XX}(t_1,t_2)\}.$$

We start with the definition

$$\begin{aligned}
K_{YY}(t_1,t_2) &\triangleq E\left[(Y(t_1)-\mu_Y(t_1))(Y^*(t_2)-\mu_Y^*(t_2))\right] \\
&= R_{YY}(t_1,t_2)-\mu_Y(t_1)\mu_Y^*(t_2) \\
&= L_{t_1}L_{t_2}\{R_{XX}(t_1,t_2)\} - L_{t_1}L_{t_2}\{\mu_X(t_1)\mu_X(t_2)\} \\
&= L_{t_1}L_{t_2}\{K_{XX}(t_1,t_2)\},
\end{aligned}$$

as was to be shown.

25. (a)

$$\dot{\mu}_Y(t) + a\mu_Y(t) = \mu_X \qquad \text{and} \qquad h(t) = e^{-at}u(t), \quad \text{so}$$

$$\begin{aligned}
\mu_Y(t) &= \mu_X \int_0^{\infty} e^{-at}dt \\
&= \mu_X/a.
\end{aligned}$$

(b) From (9.5-5a)

$$\begin{aligned}
R_{XY}(\tau) &= h^*(-\tau) * R_{XX}(\tau) \\
&= \int_{-\infty}^{\infty} h^*(-t)R_{XX}(\tau-t)dt \\
&= \int_{-\infty}^{\infty} h^*(t')R_{XX}(\tau+t')dt', \qquad \text{with } \ t' \triangleq -t, \\
&= \int_0^{\infty} e^{-at'}[\delta(\tau+t')+\mu_X^2]dt' \\
&= \begin{cases} e^{a\tau}+\mu_X^2/a, & \tau \le 0, \\ 0+\mu_X^2/a, & \tau > 0, \end{cases} \\
&= e^{a\tau}u(-\tau)+\mu_X^2/a \qquad \text{for all } \tau.
\end{aligned}$$

Then, from (9.5-5b),

$$\begin{aligned}
R_{YY}(\tau) &= \int_{-\infty}^{\infty} h(\tau_1)R_{XY}(\tau-\tau_1)d\tau_1 \\
&= \int_0^{\infty} e^{-a\tau_1}[e^{a(\tau-\tau_1)}+\mu_X^2/a]d\tau_1, \qquad \text{for} \quad \tau<0, \\
&= e^{a\tau}\int_0^{\infty} e^{-2a\tau_1}d\tau_1 + \mu_X^2/a, \\
&= e^{a\tau}/2a + (\mu_X/a)^2, \ \tau < 0.
\end{aligned}$$

For $\tau > 0$,

$$\begin{aligned}
R_{YY}(\tau) &= e^{a\tau}\int_{\tau}^{\infty} e^{-2a\tau_1}d\tau_1 + (\mu_X/a)^2, \\
&= e^{-a\tau}/2a + (\mu_X/a)^2, \tau < 0.
\end{aligned}$$

Thus for all τ, we have

$$R_{YY}(\tau) = e^{-a|\tau|}/2a + (\mu_X/a)^2 .$$

(c) Since $\mu_Y = \mu_X/a$, then for the covariance, we have

$$K_{YY}(\tau) = e^{-a|\tau|}/2a \quad \text{and} \quad \sigma_Y^2 = 1/2a.$$

26. No. For example for $t_1 = 1.0$ and $t_2 = 1.9$,

$$E[X(t_1)X(t_2)] = R_{XX}[0].$$

For $t_1 = 1.2$ and $t_2 = 1.2 + 0.9 = 2.1$,

$$E[X(t_1)X(t_2)] = R_{XX}[1] \neq R_{XX}[0].$$

Hence, $X(t)$ is not WSS.

27.

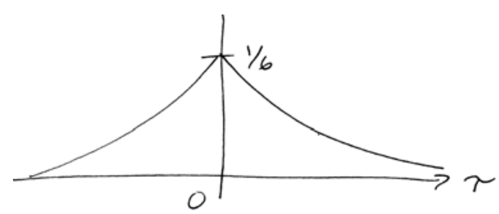
$$\begin{aligned} S_{UU}(\omega) &= \frac{1}{\omega^2+9} \\ &= \frac{1}{(j\omega+3)(-j\omega+3)} \\ &= \frac{a}{j\omega+3} + \frac{b}{-j\omega+3}, \qquad \text{using partial fraction expansion,} \\ &= \frac{1/6}{j\omega+3} + \frac{1/6}{-j\omega+3}. \end{aligned}$$

So

$$S_{UU}(s) = \frac{1/6}{s+3} + \frac{1/6}{-s+3}, \qquad |\operatorname{Re}(s)| < 3.$$

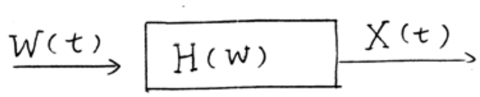
Taking the inverse Laplace transform

$$\begin{aligned} R_{UU}(\tau) &= \frac{1}{6}e^{-3\tau}u(\tau) + \frac{1}{6}e^{+3\tau}u(-\tau) \\ &= \frac{1}{6}e^{-3|\tau|}, \qquad -\infty < \tau < +\infty. \end{aligned}$$



(also see Example A.3-1 in Appendix A.)

28. Using white noise for system indentification.



$$\begin{aligned} K_{WW}(\tau) &= \delta(\tau) \\ K_{XW}(\tau) &= K_{WW}(\tau) * h(\tau) \\ &= \delta(\tau) * h(\tau) \\ &= h(\tau). \end{aligned}$$

Therefore

$$\begin{aligned} H(\omega) &= FT\{h(\tau)\} \\ &= S_{XW}(\omega). \end{aligned}$$

29. (a)

$$\begin{aligned} \mu_Y(t) &= E[X(t) + 0.3X'(t)] \\ &= \mu_X(t) + 0.3\mu'_X(t) \\ &= 5t + 0.3 \times 5 \\ &= 5t + 1.5. \end{aligned}$$

(b)

$$K_{YY}(t_1, t_2) = E[\left(X_c(t_1) + 0.3X'_c(t_1)\right)(X_c(t_2) + 0.3X'_c(t_2))],$$

where $X_c(t) \triangleq X(t) - \mu_X(t)$, i.e. the *centered version* of X. Then we can write

$$K_{YY}(t_1, t_2) = K_{XX}(t_1, t_2) + 0.3\frac{\partial K_{XX}(t_1, t_2)}{\partial t_1} + 0.3\frac{\partial K_{XX}(t_1, t_2)}{\partial t_2} + 0.09\frac{\partial^2 K_{XX}(t_1, t_2)}{\partial t_1 \partial t_2}.$$

Now, from the given K_{XX},

$$\frac{\partial K_{XX}(t_1, t_2)}{\partial t_1} = \frac{-\sigma^2 2\alpha(t_1 - t_2)}{(1 + \alpha(t_1 - t_2)^2)^2},$$

and

$$\begin{aligned} \frac{\partial K_{XX}(t_1, t_2)}{\partial t_2} &= \frac{-\sigma^2 2\alpha(t_2 - t_1)}{(1 + \alpha(t_2 - t_1)^2)^2} \\ &= -\frac{\partial K_{XX}(t_1, t_2)}{\partial t_1}. \end{aligned}$$

Thus the two cross-terms cancel, and inserting

$$\frac{\partial^2 K_{XX}(t_1, t_2)}{\partial t_1 \partial t_2} = \frac{-6\alpha^2\sigma^2 2\alpha(t_1 - t_2)^2 + 2\alpha\sigma^2}{(1 + \alpha(t_1 - t_2)^2)^3},$$

we have finally

$$K_{YY}(t_1, t_2) = \frac{\sigma^2}{1 + \alpha\tau^2}\left(1 + \frac{-0.54\alpha^2\tau^2 + 0.18\alpha}{(1 + \alpha\tau^2)^2}\right), \quad \text{with } \tau \triangleq t_1 - t_2.$$

Equivalently

$$K_{YY}(\tau) = \frac{\sigma^2}{1 + \alpha\tau^2}\left(1 + \frac{-0.54\alpha^2\tau^2 + 0.18\alpha}{(1 + \alpha\tau^2)^2}\right).$$

(c) No, not WSS. The covariance function is shift-invariant, but not the mean function.

30. (a) We have $H(\omega) = \frac{1}{j\omega+1}$ and $S_{WW}(\omega) = 1$, so

$$\begin{aligned} S_{XX}(\omega) &= |H(\omega)|^2 S_{WW}(\omega) \\ &= H(\omega)H^*(\omega)S_{WW}(\omega) \\ &= \frac{1}{j\omega+1}\frac{1}{-j\omega+1}1 \\ &= \frac{1}{\omega^2+1}. \end{aligned}$$

(b) Using the residue method, we have

$$R_{XX}(\tau) = \begin{cases} \text{Res}\left[\frac{e^{s\tau}}{(s+1)(-s+1)}; s=-1\right], & \tau \geq 0, \\ \text{Res}\left[\frac{e^{s\tau}}{(s+1)(-s+1)}; s=+1\right], & \tau \leq 0. \end{cases}$$

Now,

$$\begin{aligned} \text{Res}\left[\frac{e^{s\tau}}{(s+1)(-s+1)}; s=-1\right] &= \left\{\frac{e^{s\tau}(s+1)}{(s+1)(-s+1)}\right|_{s=-1} \\ &= \frac{1}{2}e^{-\tau}, \tau \geq 0, \end{aligned}$$

and

$$\begin{aligned} \text{Res}\left[\frac{e^{s\tau}}{(s+1)(-s+1)}; s=+1\right] &= \left\{\frac{e^{s\tau}(-s+1)}{(s+1)(-s+1)}\right|_{s=+1} \\ &= \frac{1}{2}e^{+\tau}, \tau \leq 0. \end{aligned}$$

Overall, we then have

$$R_{XX}(\tau) = \frac{1}{2}e^{-|\tau|}, -\infty < \tau < +\infty.$$

(c) Consider $X'(t) + X(t) = W(t)$ for $t > 0$, subject to initial condition $X(0)$ with $X(0) \perp W(t)$. Then, we can write

$$\begin{aligned} X(t) &= X(0)e^{-t} + \int_0^t e^{-(t-\tau)}W(\tau)d\tau, \\ &= X(0)e^{-t} + e^{-t}\int_0^t e^{\tau}W(\tau)d\tau,\ t \geq 0. \end{aligned}$$

Thus

$$\begin{aligned}
R_{XX}(t,s) &= E[X^2(0)]e^{-(t+s)} + e^{-(t+s)} \int_0^t \int_0^s e^{\tau_1+\tau_2} E[W(\tau_1)W(\tau_2)]d\tau_1 d\tau_2 \\
&= E[X^2(0)]e^{-(t+s)} + e^{-(t+s)} \int_0^t \int_0^s e^{\tau_1+\tau_2} \delta(\tau_1-\tau_2) d\tau_1 d\tau_2 \\
&= E[X^2(0)]e^{-(t+s)} + e^{-(t+s)} \int_0^s e^{\tau_2} \left(\int_0^t e^{\tau_1} \delta(\tau_1-\tau_2) d\tau_1 \right) d\tau_2 \\
&= E[X^2(0)]e^{-(t+s)} + e^{-(t+s)} \int_0^s e^{\tau_2} \left\{ \begin{array}{ll} e^{\tau_2}, & \tau_2 < t \\ 0, & \tau_2 > t \end{array} \right\} d\tau_2 \\
&= E[X^2(0)]e^{-(t+s)} + e^{-(t+s)} \int_0^{\min(t,s)} e^{2\tau_2} d\tau_2 \\
&= E[X^2(0)]e^{-(t+s)} + e^{-(t+s)} \frac{1}{2}[e^{2\min(t,s)} - 1] \\
&= E[X^2(0)]e^{-(t+s)} + \frac{1}{2}[e^{-(t-s)} - e^{-(t+s)}].
\end{aligned}$$

Choosing initial average power $E[X^2(0)] = 1/2$, we get

$$R_{XX}(t,s) = \frac{1}{2} e^{-(t-s)} \quad \text{for all } t \geq 0, s \geq 0.$$

With the initial mean $\mu_{X(0)} = E[X(0)] = 0$ also, we get finally $\mu_X(t) = E[X(t)] = 0$. Thus the desired initial conditions are

$$\mu_{X(0)} = 0 \quad \text{and} \quad \sigma^2_{X(0)} = \frac{1}{2}.$$

31. Let

$$\begin{aligned}
r(\tau) &\triangleq \int_{-\infty}^{+\infty} h^*(-u)h(\tau-u)du \\
&= h^*(-\tau) * h(\tau).
\end{aligned}$$

We need to show that

$$\sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j^* r(\tau_i - \tau_j) \geq 0, \quad \text{for arbitrary } a_i, \tau_i.$$

Consider

$$\begin{aligned}
r(\tau_i - \tau_j) &= \int_{-\infty}^{+\infty} h^*(-u)h(\tau_i - \tau_j - u)du \\
&= \int_{-\infty}^{+\infty} h(\tau_i - \tau_j - u)h^*(-u)du
\end{aligned}$$

Making the substitution $u' \triangleq -u - \tau_j$, we then get

$$r(\tau_i - \tau_j) = \int_{-\infty}^{+\infty} h(\tau_i + u')h^*(\tau_j + u')du'.$$

So now,

$$
\begin{aligned}
\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j^* r(\tau_i - \tau_j) &= \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j^* \int_{-\infty}^{+\infty} h(\tau_i + u')h^*(\tau_j + u')du' \\
&= \int_{-\infty}^{+\infty} \left(\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j^* h(\tau_i + u')h^*(\tau_j + u') \right) du' \\
&= \int_{-\infty}^{+\infty} \left| \sum_{i=1}^{N} a_i h(\tau_i + u') \right|^2 du' \\
&\geq 0, \quad \text{since integral of non-negative function.}
\end{aligned}
$$

32. (a) The input process $X(t)$ has constant mean 128. So $\mu_Y(t) = \mu_X H(0) = 128 \times 1 = 128$.

(b) For the covariance function,

$$
K_{YY}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |H(\omega)|^2 S_{X_cX_c}(\omega) e^{+j\omega\tau} d\omega, \tag{p32eq1}
$$

where $X_c \triangleq X - \mu_X$ is the centered version of X. Also

$$
\begin{aligned}
|H(\omega)|^2 &= H(\omega)H^*(\omega) \\
&= \frac{1}{1+j\omega}\frac{1}{1-j\omega} \\
&= \frac{1}{1+\omega^2}.
\end{aligned}
$$

The PSD $S_{X_cX_c}$ is determined as the FT of K_{XX} :

$$
\begin{aligned}
S_{X_cX_c}(\omega) &= \int_{-\infty}^{+\infty} 1000 e^{-10|\tau|} e^{-j\omega\tau} d\tau \\
&= 1000 \left(\int_{-\infty}^{0} e^{+(10-j\omega)\tau} d\tau + \int_{0}^{+\infty} e^{-(10+j\omega)\tau} d\tau \right) \\
&= 1000 \left(\frac{1}{10-j\omega} + \frac{1}{10+j\omega} \right) \\
&= \frac{20,000}{100+\omega^2}.
\end{aligned}
$$

Now, we can plug into (p32eq1) to obtain

$$
\begin{aligned}
K_{YY}(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{1+\omega^2} \frac{20,000}{100+\omega^2} e^{+j\omega\tau} d\omega \\
&= \frac{20,000}{2\pi(99)} \int_{-\infty}^{+\infty} \left(\frac{1}{1+\omega^2} - \frac{1}{100+\omega^2} \right) e^{+j\omega\tau} d\omega \\
&= \frac{20,000}{99} \left(\frac{1}{2} e^{-|\tau|} - \frac{1}{20} e^{-10|\tau|} \right) \\
&\doteq 101.01 e^{-|\tau|} - 10.10 e^{-10|\tau|}.
\end{aligned}
$$

33. (a) In general, we have $S_Y(\omega) = |H(\omega)|^2 S_X(\omega)$. Here $S_X(\omega) = 2$ and

$$\begin{aligned} H(\omega) &= \frac{1}{4}\int_{-2}^{+2} e^{-j\omega t} dt \\ &= \frac{e^{-jw2} - e^{+j\omega 2}}{-j2(2\omega)} \\ &= \frac{\sin 2\omega}{2\omega}. \end{aligned}$$

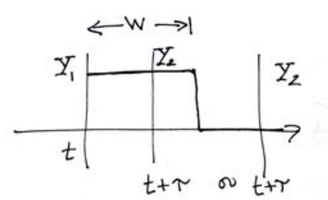
Thus

$$\begin{aligned} S_Y(\omega) &= |H(\omega)|^2 S_X(\omega) \\ &= \left(\frac{\sin 2\omega}{2\omega}\right)^2 2. \end{aligned}$$

34. We first express

$$\begin{aligned} R_{YY}(\tau) &\triangleq E[Y(t+\tau)Y(t)] \\ &= E_W[E[Y(t+\tau)Y(t)|W]], \end{aligned}$$

where the outer expectation $E_W[\cdot]$ is on the random variable W, which indicates the width of a pulse. Next, let $W = w$ and evaluate $E[Y(t+\tau)Y(t)|W = w]$. Since this quantity is not a function of t, by the WSS condition on Y, we can take t at the start of a pulse, as shown below.



Now for $\tau < w$, we have $Y_1 = Y_2$, so $Y_1 Y_2 = Y_1^2$, while for $\tau > w$, we have $Y_1 Y_2$ with $Y_1 \perp Y_2$, thus

$$\begin{aligned} E[Y(t+\tau)Y(t)|W = w] &= E[Y_1 Y_2] \\ &= \begin{cases} (E[Y])^2, & \tau > w, \\ E[Y^2], & \tau < w. \end{cases} \end{aligned}$$

Thus

$$\begin{aligned} E[Y(t+\tau)Y(t)] &= E_W[E[Y(t+\tau)Y(t)|W]] \\ &= (E[Y])^2 P[W < \tau] + E[Y^2]P[W > \tau]. \end{aligned}$$

Calculating, we get

$$\begin{aligned} P[W > \tau] &= \int_\tau^\infty \lambda e^{-\lambda w} dw \\ &= e^{-\lambda\tau}, \end{aligned}$$

and $E[Y] = 0$ and $E[Y^2] = \sigma_X^2$ since, by the problem statement, the independent pulse amplitude X has zero mean and variance σ_X^2. Combining results, we get

$$\begin{aligned} R_{YY}(\tau) &= (E[Y])^2 P[W < \tau] + E[Y^2]P[W > \tau] \\ &= 0 + E[Y^2]P[W > \tau] \\ &= \sigma_X^{22} e^{-\lambda\tau}, \qquad \tau > 0. \end{aligned}$$

Since $R_{YY}(\tau)$ must be even in the delay variable τ, we finally have

$$R_{YY}(\tau) = \sigma_X^{22} \exp(-\lambda|\tau|), \qquad \text{for all } -\infty < \tau + \infty.$$

The psd is then given as

$$S_{YY}(\omega) = \frac{2\lambda\sigma_X^2}{\omega^2 + \lambda^2}.$$

For $\tau = 0$, we get $R_{YY}(0) = \sigma_X^2$, the mean-square value of X, as it should be. For $\tau = \infty$, we get $R_{YY}(\infty) = 0$, the mean value of the process squared. Widely separated elements become uncorrelated.

35. (a)

$$S_{XX}(\omega) = 2\pi\mu_X^2\delta(\omega) + \frac{10}{(\omega - 10)^2 + 4} + \frac{10}{(\omega + 10)^2 + 4},$$

where we have used the known Fourier transform pair $\exp(-\alpha\,|\tau|) \Longleftrightarrow 2\alpha/(\omega^2 + \alpha^2)$.

(b)

$$\frac{1}{2\pi}\int_{\omega_1}^{\omega_2} (\cdot)d\omega + \frac{1}{2\pi}\int_{-\omega_2}^{-\omega_1} (\cdot)d\omega.$$

(c)

$$S_{YY}(\omega) = |H(\omega)|^2 S_{XX}(\omega),$$

with

$$\begin{aligned} |H(\omega)|^2 &= \frac{j\omega + 4}{(j\omega + 6)(j\omega + 5)} + \frac{-j\omega + 4}{(-j\omega + 6)(-j\omega + 5)} \\ &= \frac{\omega^2 + 16}{(\omega^2 + 36)(\omega^2 + 25)}. \end{aligned}$$

36. (a) Given $K_{XX}(\tau) = \frac{1}{\tau_0}e^{-|\tau|/\tau_0}$, since X is zero mean, $K_{XX} = R_{XX}$, hence

$$S_{XX}(\omega) = \frac{1}{1 + (\omega\tau_0)^2}.$$

Therefore

$$\begin{aligned} S_{YY}(\omega) &= S_{XX}(\omega)|G(\omega)|^2 \\ &= \begin{cases} \frac{1}{1+(\omega\tau_0)^2}, & |\omega| \le \omega_0, \\ 0, & |\omega| > \omega_0. \end{cases} \end{aligned}$$

(b) $S_{WW}(\omega) = FT\{K_{WW}(\tau)\} = FT\{\delta(\tau)\} = 1$, where $K(\tau) = R(\tau)$ for a zero-mean process. Thus,

$$S_{VV}(\omega) = \begin{cases} 1, & |\omega| \le \omega_0, \\ 0, & |\omega| > \omega_0. \end{cases}$$

(c) If $|\omega_0\tau_0| \ll 1$, then $1+(\omega\tau_0)^2 \approx 1$ for $|\omega| \le \omega_0$, therefore

$$\begin{aligned} S_{YY}(\omega) &\approx \begin{cases} 1, & |\omega| \le \omega_0, \\ 0, & |\omega| > \omega_0. \end{cases} \\ &= S_{VV}(\omega). \end{aligned}$$

For the correlation function error

$$\begin{aligned} R_{VV}(0) - R_{YY}(0) &= \frac{1}{2\pi}\int_{-\omega_0}^{+\omega_0} \left(1 - \frac{1}{1+(\omega\tau_0)^2}\right) d\omega \\ &= \frac{2}{2\pi}\int_0^{+\omega_0} \frac{(\omega\tau_0)^2}{1+(\omega\tau_0)^2} d\omega \\ &< \frac{2}{2\pi}\int_0^{+\omega_0} \frac{(\omega_0\tau_0)^2}{1+(\omega_0\tau_0)^2} d\omega, \quad \text{since the integrand is increasing function,} \\ &= \frac{2}{2\pi}\frac{(\omega_0\tau_0)^2}{1+(\omega_0\tau_0)^2}\int_0^{+\omega_0} d\omega \\ &= \frac{\omega_0(\omega_0\tau_0)^2}{\pi\,(1+(\omega_0\tau_0)^2)}. \end{aligned}$$

37. The processes $X(t)$ and $N(t)$ are WSS and mutually uncorrelated with zero means and psd's S_{XX} and S_{NN}, respectively.

(a) Let $Z(t) = X(t) + N(t)$, then

$$\begin{aligned} R_{ZZ}(\tau) &= E[(X(t+\tau) + N(t+\tau))(X^*(t) + N^*(t))] \\ &= R_{XX}(\tau) + R_{NN}(\tau). \end{aligned}$$

Therefore,

$$\begin{aligned} S_{YY}(\omega) &= |H(\omega)|^2 S_{ZZ}(\omega) \\ &= |H(\omega)|^2 (S_{XX}(\omega) + S_{NN}(\omega)). \end{aligned}$$

(b) For the cross-correlation,

$$\begin{aligned} R_{XY}(\tau) &= E[X(t+\tau)(h^*(t) * (X^*(t) + N^*(t))] \\ &= \int_{-\infty}^{+\infty} h^*(s) E[X(t+\tau)X^*(t-s)ds + \int_{-\infty}^{+\infty} h^*(s) E[X(t+\tau)N^*(t-s)]ds \\ &= \int_{-\infty}^{+\infty} h^*(s) R_{XX}(\tau+s)ds + \int_{-\infty}^{+\infty} h^*(s)\, 0\, ds \\ &= \int_{-\infty}^{+\infty} h^*(-s') R_{XX}(\tau - s')ds' \quad \text{with} \quad s' \triangleq -s, \\ &= h^*(-\tau) * R_{XX}(\tau). \end{aligned}$$

Therefore $S_{XY}(\omega) = H^*(\omega) S_{XX}(\omega)$. Following the same path, we can show that cross-power spectral density $S_{YX}(\omega) = H(\omega) S_{XX}(\omega)$.

(c) For the error $\xi(t) \triangleq Y(t) - X(t)$, we get

$$\begin{aligned} R_{\xi\xi}(\tau) &= E[(Y(t+\tau) - X(t+\tau))(Y^*(t) - X^*(t))] \\ &= R_{YY}(\tau) - R_{XY}(\tau) - R_{YX}(\tau) + R_{XX}(\tau). \end{aligned}$$

Proceeding to the psd's and using results from parts (a) and (b),

$$\begin{aligned} S_{\xi\xi}(\omega) &= S_{YY}(\omega) - S_{XY}(\omega) - S_{YX}(\omega) + S_{XX}(\omega) \\ &= |H(\omega)|^2(S_{XX}(\omega) + S_{NN}(\omega)) - H^*(\omega)S_{XX}(\omega) - H^*(\omega)S_{XX}(\omega) + S_{XX}(\omega) \\ &= |H(\omega) - 1|^2 S_{XX}(\omega) + |H(\omega)|^2 S_{NN}(\omega). \end{aligned}$$

(d) Given $h(t) = a\delta(t)$, then $H(\omega) = a$. Letting a be real, we then have

$$S_{\xi\xi}(\omega) = (a^2 - 2a + 1)S_{XX}(\omega) + a^2 S_{NN}(\omega),$$

or equivalently

$$R_{\xi\xi}(\tau) = (a^2 - 2a + 1)R_{XX}(\tau) + a^2 R_{NN}(\tau).$$

To find the minimum of $R_{\xi\xi}(0)$ with respect to the variable a, we differentiate and set the derivative to zero, obtaining

$$\begin{aligned} \frac{dR_{\xi\xi}(0)}{da} &= (2a - 2)R_{XX}(0) + 2aR_{NN}(0) \\ &= 0. \end{aligned}$$

Thus

$$a_{\min} = \frac{R_{XX}(0)}{R_{XX}(0) + R_{NN}(0)}.$$

To see that this is a true minimum, and not merely a stationary point, we calculate the second derivative and find

$$\begin{aligned} \frac{d^2 R_{\xi\xi}(0)}{da^2} &= 2R_{XX}(0) + 2R_{NN}(0) \\ &\geq 0 \quad \text{always for valid correlation functions, and} \\ &> 0 \quad \text{if } X(t) \text{ has positive average power.} \end{aligned}$$

38.

39.

40. (a) See definition 9.2-1.

(b) See definition 9.2-4.

(c)

$$\begin{aligned} f(x_{t_n}|x_{t_{n-1}}, x_{t_{n-2}}, ..., x_{t_1}) &= f(x_{t_n} - x_{t_{n-1}}|x_{t_{n-1}}, x_{t_{n-2}}, ..., x_{t_1}), \quad \text{since } x_{t_{n-1}} \text{ is conditionally known,} \\ &= f(x_{t_n} - x_{t_{n-1}}), \quad \text{by independent increments property,} \\ &= f(x_{t_n}|x_{t_{n-1}}), \quad \text{by Markov property.} \end{aligned}$$

41. $Y(t) = X(t) - X(T)$ for $t \geq T$. So $\mu_Y(t) = \mu_X(t) - \mu_X(T) = \mu_0 - \mu_0 = 0$, for $t \geq T$. Next consider the covariance function of $Y(t)$.

$$\begin{aligned} K_{YY}(t_1, t_2) &= E[Y(t_1)Y(t_2)] \\ &= E[(X(t_1) - X(T))\,(X(t_2) - X(T))]. \end{aligned}$$

Consider the case $t_2 > t_1$, then we can write the second factor $(X(t_2) - X(T))$ as

$$(X(t_2) - X(t_1)) + (X(t_1) - X(T))\,,$$

with both terms independent because of the independent increments property. So for the above expectation, we have

$$\begin{aligned} E[(X(t_1) - X(T))\,(X(t_2) - X(T))] &= E[(X(t_1) - X(T))\,(X(t_2) - X(t_1))] \\ &\quad + E[(X(t_1) - X(T))\,(X(t_1) - X(T))] \\ &= E[X(t_1) - X(T)]E[X(t_2) - X(t_1)] + E[(X(t_1) - X(T))^2] \\ &= 0 \times 0 + E[(X(t_1) - X(T))^2] \\ &= E[(X(t_1) - X(T))^2], \quad \text{for all } T \leq t_1 < t_2 \end{aligned}$$

So, we now have to determine $E[(X(t_1) - X(T))^2]$. Exploiting independent increments again, we can write

$$X(t_1) = (X(t_1) - X(T)) + X(T),$$

where the two terms are independent, since $t_1 \geq T$. Thus variances add, and we have

$$\sigma_X^2(t_1) = E[(X(t_1) - X(T))^2] + \sigma_X^2(T),$$

or $E[(X(t_1) - X(T))^2] = \sigma_X^2(t_1) - \sigma_X^2(T)$. So for $t_2 > t_1 (\geq T)$, we have $K_Y(t_1, t_2) = \sigma_X^2(t_1) - \sigma_X^2(T)$. By the symmetry of covariance function, it must be that $K_Y(t_1, t_2) = \sigma_X^2(t_2) - \sigma_X^2(T)$ when $t_2 < t_1$. Note that at $t_1 = t_2$ both answers would be valid. Thus the full answer is

$$K_{YY}(t_1, t_2) = \sigma_X^2\left(\min(t_1, t_2) - \sigma_X^2(T)\right), \quad \text{for all } t_1, t_2 \geq T.$$

42.

43. We are given the third-order random (stochastic) differential equation

$$\frac{d^3Y(t)}{d^3t} + a_2\frac{d^2Y(t)}{d^2t} + a_1\frac{dY(t)}{dt} + a_0Y(t) = X(t).$$

(a) Putting this into vector form, we have with $\mathbf{Y}(t) \triangleq (Y(t), \dot{Y}(t), \ddot{Y}(t))^T$,

$$\dot{\mathbf{Y}}(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \mathbf{Y}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} X(t), \qquad ((1))$$

then

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

(b) We have

$$\dot{\mathbf{Y}}^\dagger(t) = \mathbf{Y}^\dagger(t)\mathbf{A}^\dagger + X(t)\mathbf{B}^\dagger. \qquad ((2))$$

Turning to correlations, then

$$\begin{aligned}
\mathbf{R}_{X\dot{\mathbf{Y}}}(\tau) &= E[X(t+\tau)\dot{\mathbf{Y}}^\dagger(t)] \\
&= E[X(t+\tau)\mathbf{Y}^\dagger(t)\mathbf{A}^\dagger] + E[X(t+\tau)X(t)\mathbf{B}^\dagger] \\
&= \mathbf{R}_{X\mathbf{Y}}(\tau)\mathbf{A}^\dagger + R_{XX}(\tau)\mathbf{B}^\dagger \\
&= -\frac{d\mathbf{R}_{X\mathbf{Y}}(\tau)}{d\tau}.
\end{aligned}$$

(c) Using above equation (1)

$$\begin{aligned}
E[\dot{\mathbf{Y}}(t+\tau)\mathbf{Y}^\dagger(t)] &= E[\mathbf{A}\mathbf{Y}(t+\tau)\mathbf{Y}^\dagger(t)] + E[\mathbf{B}X(t+\tau)\mathbf{Y}^\dagger(t)] \\
&= \mathbf{A}\mathbf{R}_{\mathbf{YY}}(\tau) + \mathbf{B}\mathbf{R}_{X\mathbf{Y}}(\tau) \\
&= \frac{d\mathbf{R}_{\mathbf{YY}}(\tau)}{d\tau}.
\end{aligned}$$

(d) Taking the Fourier transforms of the results in part (b), we get

$$-j\omega\mathbf{S}_{X\mathbf{Y}}(\omega) = \mathbf{S}_{X\mathbf{Y}}(\omega)\mathbf{A}^\dagger + S_{XX}(\omega)\mathbf{B}^\dagger.$$

Solving for $\mathbf{S}_{X\mathbf{Y}}$, we then obtain

$$\mathbf{S}_{X\mathbf{Y}}(\omega) = S_{XX}(\omega)\mathbf{B}^\dagger(-j\omega\mathbf{I} - \mathbf{A}^\dagger)^{-1}.$$

From part (c), taking Fourier transforms, we obtain $(j\omega\mathbf{I} - \mathbf{A})\mathbf{S}_{\mathbf{YY}}(\omega) = \mathbf{B}\mathbf{S}_{X\mathbf{Y}}(\omega)$,with solution

$$\mathbf{S}_{\mathbf{YY}}(\omega) = (j\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{S}_{X\mathbf{Y}}(\omega).$$

Combining these two results, we finally get the matrix psd

$$\mathbf{S}_{\mathbf{YY}}(\omega) = (j\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}S_{XX}(\omega)\mathbf{B}^\dagger(-j\omega\mathbf{I} - \mathbf{A}^\dagger)^{-1}.$$

44.

45. (a) $Y(t) = \mathbf{A}^T\mathbf{X}(t)$, so

$$\begin{aligned}
R_{YY}(\tau) &= E[Y(t+\tau)Y^*(t)] \\
&= E[\mathbf{A}^T\mathbf{X}(t+\tau)\mathbf{X}^\dagger(t)\mathbf{A}^*] \\
&= \mathbf{A}^T\mathbf{R}_{\mathbf{XX}}(\tau)\mathbf{A}^*.
\end{aligned}$$

Thus

$$\begin{aligned}
S_{YY}(\omega) &\triangleq FT\{R_{YY}(\tau)\} \\
&= FT\{\mathbf{A}^T\mathbf{R}_{\mathbf{XX}}(\tau)\mathbf{A}^*\} \\
&= \sum_{i,k} a_i a_k^* FT\{(\mathbf{R}_{\mathbf{XX}}(\tau))_{i,k} \\
&= \mathbf{A}^T\mathbf{S}_{\mathbf{XX}}(\omega)\mathbf{A}^*.
\end{aligned}$$

(b) Since $S_{YY}(\omega) \geq 0$ for all ω and for all vectors $\mathbf{A}$, then part (a) shows that $\mathbf{S}_{\mathbf{XX}}(\omega)$ must be a positive semidefinite matrix for each frequency ω.

46.

47.

48.

(a)

$$\begin{aligned} P_1(t+dt) &= P_1(t)(1-1dt) + P_2(t)(2dt) \\ P_2(t+dt) &= P_1(t)(1dt) + P_2(t)(1-2dt), \end{aligned}$$

or, forming derivative terms on the left-hand side,

$$\begin{aligned} \frac{P_1(t+dt)-P_1(t)}{dt} &= -P_1(t) + 2P_2(t) \\ \frac{P_2(t+dt)-P_2(t)}{dt} &= P_1(t) - 2P_2(t), \end{aligned}$$

(b) For steady state, $dP_i/dt = 0$, which then implies $P_1 = 2P_2$, but always $P_1 + P_2 = 1$, so we have asympotically, the steady-state probabilities:

$$P_1 = \frac{2}{3} \quad \text{and} \quad P_2 = \frac{1}{3}.$$

49. Since X and Y are jointly WSS,

$$R_{XY}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_{XY}(\omega)e^{+j\omega\tau}d\omega,$$

then directly from part (a), we get

$$\left|\int_{-\infty}^{+\infty} S_{XY}(\omega)e^{+j\omega\tau}d\omega\right| \leq \sqrt{\int_{-\infty}^{+\infty} S_{XX}(\omega)d\omega \int_{-\infty}^{+\infty} S_{YY}(\omega)d\omega}.$$

Then, putting $X(t)$ and $Y(t)$ through narrow-band filters with gain 1, bandwidth ϵ (a very small positive number), and centered at frequency ω_0, call the respective outputs $\widetilde{X}(t)$ and $\widetilde{Y}(t)$ Then, written for processes $\widetilde{X}(t)$ and $\widetilde{Y}(t)$, this above inequality becomes

$$\begin{aligned} \left|\int_{-\infty}^{+\infty} S_{\widetilde{X}\widetilde{Y}}(\omega)e^{+j\omega\tau}d\omega\right| &\approx S_{XY}(\omega_0)|e^{+j\omega_0\tau}|\epsilon \\ &= S_{XY}(\omega_0)\epsilon \\ &\leq \sqrt{S_{XX}(\omega_0)\epsilon\, S_{YY}(\omega_0)\epsilon} \\ &= \sqrt{S_{XX}(\omega_0)S_{YY}(\omega_0)}\epsilon. \end{aligned}$$

Then, cancelling out ϵ and letting $\epsilon \searrow 0$, we get the exact inequality

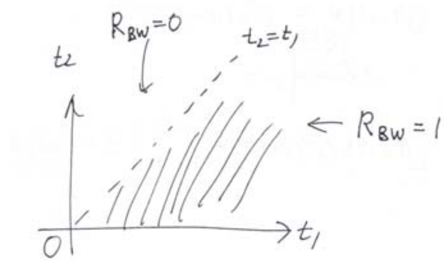
$$S_{XY}(\omega_0) \leq \sqrt{S_{XX}(\omega_0)S_{YY}(\omega_0)}.$$

Finally since ω_0 is an arbitrary frequency, this inequality must be true for all $-\infty < \omega < +\infty$, as was to be shown.

50. (a)

$$
\begin{aligned}
R_{BW}(t_1,t_2) &= E\left[\int_0^{t_1} W(\tau_1)d\tau_1 W(t_2)\right] \\
&= \int_0^{t_1} R_{WW}(\tau_1,t_2)d\tau_1 \\
&= \int_0^{t_1} \delta(\tau_1 - t_2)d\tau_1 \\
&= u(t_1 - t_2).
\end{aligned}
$$

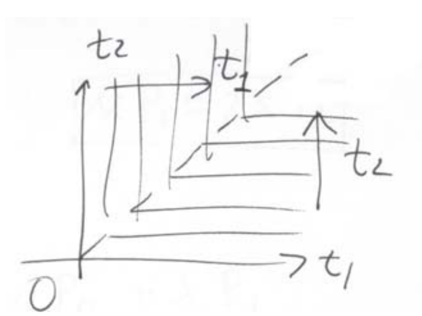
Here is a sketch.



(b) Let the two times $t_1, t_2 \geq 0$, then

$$
\begin{aligned}
R_{BB}(t_1,t_2) &= E\left[\int_0^{t_1}\int_0^{t_2} W(\tau_1)W(\tau_2)d\tau_1 d\tau_2\right] \\
&= \int_0^{t_1}\int_0^{t_2} R_{WW}(\tau_1,\tau_2)d\tau_1 d\tau_2 \\
&= \int_0^{t_1}\int_0^{t_2} \delta(\tau_1-\tau_2)d\tau_1 d\tau_2 \\
&= \int_0^{t_1}\left(\int_0^{t_2} \delta(\tau_1-\tau_2)d\tau_2\right)d\tau_1 \\
&= \int_0^{t_1} u(t_2-\tau_1)d\tau_1 \\
&= \min(t_1,t_2), t_1,t_2 \geq 0.
\end{aligned}
$$

Here is a sketch.



51. (a)

$$\begin{aligned}
p_0(t+\delta t) &= (1-2\mu\,\delta t)p_0(t) + \lambda\,\delta t\,p_1(t) + 0\\
p_1(t+\delta t) &= 2\mu p_0(t) + (1-(\mu+\lambda)\delta t)p_1(t) + 2\lambda\,\delta t\,p_2(t)\\
p_2(t+\delta t) &= 0 + \mu\,\delta t\,p_1(t) + (1-2\lambda\delta t)p_2(t),
\end{aligned}$$

which yeilds $d\mathbf{p}(t)/dt = \mathbf{A}\mathbf{p}(t)$ with

$$\mathbf{A} \triangleq \begin{bmatrix} -2\mu & +\lambda & 0 \\ +2\mu & -(\mu+\lambda) & +2\lambda \\ 0 & +\mu & -2\lambda \end{bmatrix}.$$

(b) In steady-state, the state probabilities $p_i(t)$ are constant, so the flow or rate of increase over δt must balance out between adjacent states $(i-1)$ and i. If all the net flows are zero, then $p_i(t)$ cannot change. Looking at the state-transition diagram, we see the flow from state 0 to state 1 is $2\mu p_0(t)$, while the flow from state 1 to state 0 is $\lambda p_1(t)$, so to balance, we must have $2\mu p_0 = \lambda p_1$ in steady state. Between states 1 and 2, we similarly balance the flow by setting $\mu p_1 = 2\lambda p_2$. To see whether this solution solves $\mathbf{Ap} = \mathbf{0}$, we write its resulting three equations

$$\begin{aligned}
-2\mu p_0 + \lambda p_1 \qquad &= 0,\\
+2\mu p_0 - (\mu+\lambda)p_1 + 2\lambda p_2 &= 0,\\
+\mu p_1 - 2\lambda p_2 &= 0.
\end{aligned}$$

We can see that the two flow solutions are equivalent to the first and third equation arising from $\mathbf{Ap} = \mathbf{0}$. Since the middle equation is obtained by summing the first and last one, and then multiplying by -1, we conclude that $\mathbf{Ap} = \mathbf{0}$ is satisfied.

(c) From our two flow equations

$$p_1 = \frac{2\mu}{\lambda}p_0 \quad \text{and} \quad p_2 = \frac{\mu}{2\lambda}p_1 = \left(\frac{\mu}{\lambda}\right)^2 p_0.$$

But also,

$$\begin{aligned}
1 &= p_0 + p_1 + p_2\\
&= p_0\left(1 + \frac{2\mu}{\lambda} + \left(\frac{\mu}{\lambda}\right)^2\right).
\end{aligned}$$

Thus, we have the steady-state solution

$$\begin{aligned}
p_0 &= \frac{1}{1+\frac{2\mu}{\lambda}+\left(\frac{\mu}{\lambda}\right)^2} = \frac{\lambda^2}{\lambda^2+2\mu\lambda+\mu^2},\\
p_1 &= \frac{2\mu\lambda}{\lambda^2+2\mu\lambda+\mu^2},\\
p_2 &= \frac{\mu^2}{\lambda^2+2\mu\lambda+\mu^2}.
\end{aligned}$$

For $\lambda = 0.001$ and $\mu = 0.1$, we get $\lambda^2 + 2\mu\lambda + \mu^2 \approx 10^{-2}$, and then

$$\begin{aligned}
p_0 &\approx \frac{10^{-6}}{10^{-2}} = 10^{-4},\\
p_1 &\approx \frac{2\times 10^{-4}}{10^{-2}} = 2\times 10^{-2}, \quad \text{and}\\
p_2 &\approx 1 - 2\times 10^{-2} = 0.98.
\end{aligned}$$

52. (a) We have $Y_1 = h * X_1 + g * X_2$, thus

$$\begin{aligned}
S_{Y_1X_1}(\omega) &= FT\{R_{Y_1X_1}(\tau)\} \\
&= FT\{E[(h * X_1)(t+\tau)X_1^*(t)] + E[(g * X_2)(t+\tau)X_1^*(t)]\} \\
&= FT\{h(\tau) * R_{X_1X_1}(\tau) + g(\tau) * R_{X_2X_1}(\tau)\} \\
&= FT\{h(\tau) * R_{X_1X_1}(\tau)\}, \quad \text{since } X_1 \perp X_2, \\
&= H(\omega)S_{X_1X_1}(\omega).
\end{aligned}$$

(b)

$$\begin{aligned}
S_{Y_2X_2}(\omega) &= FT\{E[Y_2(t+\tau)X_2^*(t)]\} \\
&= FT\{b(\tau) * R_{X_2X_2}(\tau)\}, \quad \text{since } \ U \perp X_2, \\
&= B(\omega)S_{X_2X_2}(\omega).
\end{aligned}$$

(c)

$$\begin{aligned}
E[Y_1(t+\tau)Y_2^*(t)] &= E[(h * X_1 + g * X_2)(t+\tau)(U^* + b^* * X_2^*)(t)] \\
&= E[(g * X_2)(t+\tau)(b^* * X_2^*)(t)], \quad \text{since } \ X_1 \perp X_2, X_1 \perp U, \text{ and } X_2 \perp U, \\
&= g(\tau) * b^*(-\tau) * R_{X_2X_2}(\tau).
\end{aligned}$$

Thus

$$\begin{aligned}
S_{Y_1Y_2}(\omega) &= FT\{g(\tau) * b^*(-\tau) * R_{X_2X_2}(\tau)\} \\
&= G(\omega)\, B^*(\omega)S_{X_2X_2}(\omega).
\end{aligned}$$

53. (a)

$$\begin{aligned}
E[Y(t_1)Y^*(t_2)] &= E\left[\sum_{n_1,n_2}^{N-1,N-1} A_{n_1}X(t_1 - n_1T)A_{n_2}^*X^*(t_2 - n_2T)\right] \\
&= \sum_{n_1,n_2}^{N-1,N-1} E[A_{n_1}A_{n_2}^*]E[X(t_1 - n_1T)X^*(t_2 - n_2T)] \\
&= \sum_{n_1,n_2}^{N-1,N-1} R_{AA}(n_1,n_2)R_{XX}(t_1 - t_2 - (n_1 - n_2)T) \\
&= R_{YY}(t_1,t_2).
\end{aligned}$$

(b) From part (a), we can see

$$\begin{aligned}
R_{YY}(t+\tau,t) &= \sum_{n_1,n_2}^{N-1,N-1} R_{AA}(n_1,n_2)R_{XX}(\tau - (n_1 - n_2)T) \\
&= R_{YY}(\tau), \quad \text{independent of } t.
\end{aligned}$$

Also

$$\mu_Y(t) = \left(\sum_n \mu_A[n]\right)\mu_X = \text{ constant wrt } t.$$

So, we don't need A_n to be a WSS random sequence, in order to make $Y(t)$ be a WSS random process.

(c) No. If A_n and $X(t)$ are correlated, we only have $E[A_n X(t-nT)] = E[A_n]E[X(t-nT)]$, not the four term product given in part (a). This is a type of 4th order moment.

54.

(a) Equating probability flows across a cut between the two states, we get $P_G \lambda_{GB} \delta t = P_B \lambda_{BG} \delta t$, so

$$\begin{aligned} P_B &= \frac{\lambda_{GB}}{\lambda_{BG}} P_G \\ &= \frac{\lambda_{GB}}{\lambda_{BG}}(1 - P_B), \end{aligned}$$

which solves to

$$P_B = \frac{\lambda_{GB}}{\lambda_{GB} + \lambda_{BG}} \qquad \text{and then,} \quad P_G = \frac{\lambda_{BG}}{\lambda_{GB} + \lambda_{BG}}.$$

(b) In the *bad* state, the transition probability time τ is an exponential RV with pdf

$$f_\tau(t) = \lambda_{BG} e^{-\lambda_{BG} t} u(t).$$

The average value is then $\mu_\tau = 1/\lambda_{BG}$, the average error-burst length.

55. Here $\lambda = 3$.

(a) For $n = 2$ and $t = 4$, we have

$$\begin{aligned} P_N(2;4) &\triangleq P[N(4) = 2] \\ &= \frac{(3 \cdot 4)^2}{2!} e^{-3 \cdot 4} \\ &= \frac{144}{2} e^{-12} \\ &= 72 e^{-12} \\ &\doteq 4.42 \times 10^{-4}. \end{aligned}$$

(b)

$$\begin{aligned} P_N(1,2;1,2) &\triangleq P[N(1) = 1, N(2) = 2] \\ &= P[N(1) = 1] P[N(1) = 1] \\ &= P_N(1;1) P_N(1;1) \\ &= P_N^2(1;1) \\ &= \left(\frac{3^1}{1!} e^{-3 \cdot 1} \right)^2 \\ &= 9 e^{-6} \\ &\doteq 0.0223. \end{aligned}$$

56. (a) $S_{YY}(\omega) = |H(\omega)|^2 (S_{XX}(\omega) + S_{VV}(\omega))$.

(b) $Y[n] = \sum_k h[k][X[n-k] + V[n-k]]$, so

$$\begin{aligned} E[X[n+m]Y^*[n]] &= \sum_k h^*[k]E[X[n+m]X^*[n-k]], \quad \text{or} \\ R_{XY}[m] &= \sum_k h^*[k]R_{XX}[m+k] \\ &= h^*[-m] * R_{XX}[m], \end{aligned}$$

so $S_{YY}(\omega) = H^*(\omega)S_{XX}(\omega)$.

57.

$$\begin{aligned} E[|\widetilde{X}(t)|^2] &= E[|X(t) + U(t)|^2] \\ &= E[|X(t)|^2] + E[|U(t)|^2] \\ &= P + \epsilon \\ &= E[\widetilde{Y}(t)|^2]. \end{aligned}$$

$$\begin{aligned} E[\widetilde{X}(t_1)\widetilde{Y}^*(t_2)] &= E[X(t_1)(Y^*(t_2) + V^*(t_2))] + E[U(t_1)(Y^*(t_2) + V^*(t_2))] \\ &= E[X(t_1)Y^*(t_2)] \\ &= \rho_{XY}(t_1, t_2)P. \end{aligned}$$

So $\rho_{\widetilde{X}\widetilde{Y}}(t_1, t_2) = \rho_{XY}(t_1, t_2)\frac{P}{P+\epsilon}$.

58.

(a)

$$\begin{aligned} & S_{YY}\left(\omega\right) = |H(\omega)|^2(S_{XX}(\omega) + S_{VV}(\omega)) \\ = \;& 100w(\omega)\left(\frac{1}{\omega^2+5} + \frac{2\omega^2+8}{(\omega^2+3)(\omega^2+5)}\right), \qquad |\omega| \le \pi, \\ = \;& 100\left(\frac{3\omega^2+11}{(\omega^2+3)(\omega^2+5)}\right)w(\omega), \end{aligned}$$

where $w(\omega) \triangleq u(\omega + \frac{\pi}{2}) - u(\omega - \frac{\pi}{2})$.

(b) Call $W[n] \triangleq X[n] + V[n]$, then

$$\begin{aligned} R_{WW}[m] &= E((X[n+m] + V[n+m])(X^*[n] + V^*[n])) \\ &= R_{XX}[m] + R_{XV}[m] + R_{VX}[m] + R_{VV}[m]. \end{aligned}$$

So, upon Fourier transformation,

$$\begin{aligned} S_{WW}(\omega) &= S_{XX}(\omega) + S_{XV}(\omega) + S_{VX}(\omega) + S_{VV}(\omega) \\ &= S_{XX}(\omega) + S_{XV}(\omega) + S_{XV}(\omega) + S_{VV}(\omega), \end{aligned}$$

since S_{VX} is real valued. Proceeding, we get

$$\begin{aligned} S_{WW}(\omega) &= \frac{3}{\omega^2+5} + \frac{2\omega^2+8}{(\omega^2+3)(\omega^2+5)} \\ &= \frac{5\omega^2+17}{(\omega^2+3)(\omega^2+5)}. \end{aligned}$$

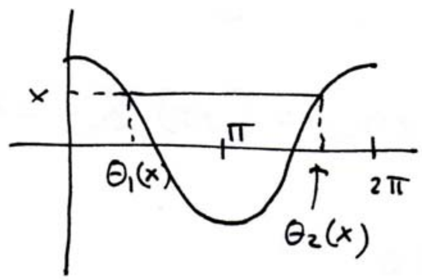
So,

$$\begin{aligned} S_{YY}(\omega) &= |H(\omega)|^2 S_{WW}(\omega) \\ &= 100\frac{5\omega^2+17}{(\omega^2+3)(\omega^2+5)}w(\omega), \qquad |\omega| \leq \pi. \end{aligned}$$

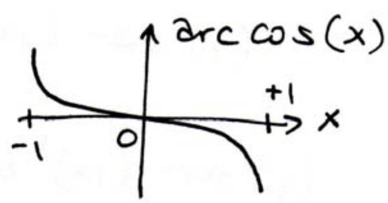
59. The easy way is to first define $\Theta_t \triangleq \omega_0 t + \Theta$, and then note that since $\Theta : U[0, 2\pi]$, so must Θ_t also be uniformly distributed on $[0, 2\pi]$. Then for fixed t, the transformation $X(t) = \cos\Theta_t = x$ has two roots θ_1 and θ_2, given as

$$\theta_1(x) = \cos^{-1} x \quad \text{and} \quad \theta_2(x) = 2\pi - \cos^{-1} x \ ,$$

as indicated in the figure below.



Here $\cos^{-1}(\cdot)$ is the principal value arc cosine as indicated in the figure below.



We then have

$$\begin{aligned} \frac{d\theta_1}{dx} &= \frac{d\cos^{-1}x}{dx} = \frac{1}{\sqrt{1-x^2}}, \quad \text{and} \\ \frac{d\theta_2}{dx} &= -\frac{d\cos^{-1}x}{dx} = \frac{-1}{\sqrt{1-x^2}}, \end{aligned}$$

so

$$\begin{aligned} f_X(x;t) &= f_\Theta(\theta_1(x))\left|\frac{d\theta_1}{dx}\right| + f_\Theta(\theta_2(x))\left|\frac{d\theta_2}{dx}\right| \\ &= \frac{1}{2\pi}\frac{1}{\sqrt{1-x^2}} + \frac{1}{2\pi}\frac{1}{\sqrt{1-x^2}}, \qquad |x| < 1, \\ &= \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}(u(x+1) - u(x-1)), \end{aligned}$$

where $u(\cdot)$ denotes the continuous parameter unit step function. Since the pdf $f_X(x;t)$ is independent of t, the process is first-order stationary by definition. The direct method is to work directly with Θ rather than defining Θ_t first. In that case, we would get

$$\theta_1(x) = \cos^{-1} x - \omega_0 t \quad \text{and} \quad \theta_2(x) = 2\pi - \cos^{-1} x - \omega_0 t \ ,$$

with Θ uniformly distributed on $[0, 2\pi]$. The problem then proceeds as above, with the derivatives unchanged. The resulting answer is the same $f_X(x;t)$ that is independent of time t.

To find the conditional density $f_X(x_2|x_1;t_2,t_1)$ for $t_2 > t_1$, we proceed as follows. Given $X(t_1) = x_1 = \cos(\omega_0 t_1 + \theta)$, then the two corresponding roots are $\theta_1 = \cos^{-1} x_1 - \omega_0 t_1$ and $\theta_2 = 2\pi - \cos^{-1} x_1 - \omega_0 t_1$ equally likely in Θ. So $X(t_2) = \cos(\omega_0 t_2 + \cos^{-1} x_1 - \omega_0 t_1)$ or $X(t_2) = \cos(\omega_0 t_2 + 2\pi - \cos^{-1} x_1 - \omega_0 t_1)$, equally likely with probability 1/2. This simplifies to

$$X(t_2) = \cos(\omega_0 (t_2 - t_1) \pm \cos^{-1} x_1), \quad \text{equally likely.}$$

Thus we end up with conditional probability

$$f_X(x_2|x_1;t_2,t_1) = \frac{1}{2}\delta(x_2 - \cos(\omega_0(t_2-t_1) + \cos^{-1} x_1)) + \frac{1}{2}\delta(x_2 - \cos(\omega_0(t_2-t_1) - \cos^{-1} x_1)).$$

Using the trigonometric identities $\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$, we can also write

$$\cos(\omega_0(t_2-t_1) \pm \cos^{-1} x_1) = x_1 \cos \omega_0 (t_2 - t_1) \mp \sqrt{1 - x^2} \sin \omega_0 (t_2 - t_1),$$

with a consequent change in the expression for $f_X(x_2|x_1;t_2,t_1)$ above.

60.

$$\begin{aligned} U(t) &\triangleq \operatorname{Re}[Z(t) e^{-j\omega_0 t}] \\ &= X(t) \cos \omega_0 t + Y(t) \sin \omega_0 t, \end{aligned}$$

since $Z = X + jY$. Then

$$\begin{aligned} E[U(t+\tau)U(t)] &= \\ & \quad E[(X(t+\tau)\cos\omega_0(t+\tau) + Y(t+\tau)\sin\omega_0(t+\tau))(X(t)\cos\omega_0 t + Y(t)\sin\omega_0 t)] \\ &= R_{XX}(\tau)\cos\omega_0(t+\tau)\cos\omega_0 t + R_{YY}(\tau)\sin\omega_0(t+\tau)\sin\omega_0 t \\ &= R_{XX}(\tau)[\cos\omega_0(t+\tau)\cos\omega_0 t + \sin\omega_0(t+\tau)\sin\omega_0 t], \qquad \text{if} \quad R_{XX}(\tau) = R_{YY}(\tau), \\ &= R_{XX}(\tau)\cos\omega_0\tau, \end{aligned}$$

after making use of the trig identity for $\cos(A - B)$. Since the mean function $\mu_U(t) = 0$, we can then say that the random process $U(t)$ is WSS. The general condition is then $R_{XX}(\tau) = R_{YY}(\tau)$.

61. By equating probability flows, we get the equalities

$$\begin{aligned} \lambda_1 P_1 &= \mu_2 P_2, \\ \lambda_2 P_2 &= \mu_3 P_3, \quad \text{and} \\ \lambda_3 P_3 &= \mu_4 P_4. \end{aligned}$$

From the first equation, $P_2 = \frac{\lambda_1}{\mu_2} P_1$, and then

$$\begin{aligned} P_3 &= \frac{\lambda_2}{\mu_3} P_2 \\ &= \frac{\lambda_2}{\mu_3} \frac{\lambda_1}{\mu_2} P_1, \end{aligned}$$

and

$$\begin{aligned} P_4 &= \frac{\lambda_3}{\mu_4} P_3 \\ &= \frac{\lambda_3}{\mu_4} \frac{\lambda_2}{\mu_3} \frac{\lambda_1}{\mu_2} P_1. \end{aligned}$$

Using the fact that these four probabilities must also sum to one, i.e. $\sum_i P_i = 1$, we finally get

$$\begin{aligned} P_1 &= \frac{1}{1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2} + \frac{\lambda_3}{\mu_4}\frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}, \\ P_2 &= \frac{\frac{\lambda_1}{\mu_2}}{1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2} + \frac{\lambda_3}{\mu_4}\frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}, \\ P_3 &= \frac{\frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}{1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2} + \frac{\lambda_3}{\mu_4}\frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}, \text{ and} \\ P_4 &= \frac{\frac{\lambda_3}{\mu_4}\frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}{1 + \frac{\lambda_1}{\mu_2} + \frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2} + \frac{\lambda_3}{\mu_4}\frac{\lambda_2}{\mu_3}\frac{\lambda_1}{\mu_2}}. \end{aligned}$$

62. (a) The probability of leaving state 2 for the first time at time t is zero, since the waiting time is an exponential RV, a continuous RV.

(b)

$$\begin{aligned} P_1(t+\delta t) &= (1 - \lambda_1\, \delta t) P_1(t) + \mu_2\, \delta t\; P_2(t) + 0 P_3(t) \\ P_2(t+\delta t) &= +\lambda_1\, \delta t\; P_1(t) - (\lambda_2 + \mu_2)\delta t\; P_2(t) + \mu_3\, \delta t\; P_3(t) \\ P_3(t+\delta t) &= 0 P_1(t) + \lambda_2\, \delta t\; P_2(t) + -\mu_3\, \delta t\; P_3(t), \end{aligned}$$

or

$$\begin{aligned} \dot{\mathbf{P}}(t) &= \underbrace{\begin{bmatrix} -\lambda_1 & +\mu_2 & 0 \\ +\lambda_1 & -(\lambda_2+\mu_2) & +\mu_3 \\ 0 & +\lambda_2 & -\mu_3 \end{bmatrix}}_{\triangleq \mathbf{A}} \mathbf{P}(t) \\ &= \mathbf{A}\mathbf{P}(t). \end{aligned}$$

(c) We substitute $\exp(\mathbf{A}t) \cdot \mathbf{P}(0)$ into this equation, and then take the term-by-term derivative of the matrix-exponential series, to obtain

$$
\begin{aligned}
\dot{\mathbf{P}}(t) &= \frac{d}{dt}\left(\sum_{k=0}^{\infty}\frac{1}{k!}(\mathbf{A}t)^k\right)\mathbf{P}(0) \\
&= \left(\sum_{k=1}^{\infty}\frac{1}{k!}k\mathbf{A}^k t^{k-1}\right)\mathbf{P}(0) \\
&= \left(\mathbf{A}\sum_{k'=0}^{\infty}\frac{1}{k'!}\,(\mathbf{A}t)^{k'}\right)\mathbf{P}(0), \quad \text{with} \quad k' \triangleq k-1, \\
&= \mathbf{A}\exp(\mathbf{A}t)\cdot\mathbf{P}(0) \\
&= \mathbf{A}\mathbf{P}(t),
\end{aligned}
$$

as was to be shown.

Chapter 10 solutions

1. (a) It was shown in the text that the m.s. limit is a linear operator, namely if $X[n] \to X$ and $Y[n] \to Y$ in the m.s.-sense, then

$$\lim_{n\to\infty}(aX[n]+bY[n]) = aX+bY. \quad \text{(m.s.)} \tag{1}$$

We are here interested in extending this relationship to derivatives, by showing:

$$\begin{aligned}
\frac{d}{dt}[aX_1(t)+bX_2(t)] &= \lim_{\epsilon\to 0}\left[\frac{(aX_1(t+\epsilon)+bX_2(t+\epsilon))-(aX_1(t)+bX_2(t))}{\epsilon}\right] \\
&= \lim_{\epsilon\to 0}\left[a\left(\frac{X_1(t+\epsilon)-X_1(t)}{\epsilon}\right)+b\left(\frac{X_2(t+\epsilon)-X_2(t)}{\epsilon}\right)\right] \\
&= a\lim_{\epsilon\to 0}\left(\frac{X_1(t+\epsilon)-X_1(t)}{\epsilon}\right)+b\lim_{\epsilon\to 0}\left(\frac{X_2(t+\epsilon)-X_2(t)}{\epsilon}\right), \quad \text{by (1)}, \\
&= a\frac{d}{dt}X_1(t)+b\frac{d}{dt}X_2(t).
\end{aligned}$$

(b) By definition, we have

$$E[X_1(t)X_2'(t)] = E\left[X_1(t)\lim_{\epsilon\to 0}\left(\frac{X_2(t+\epsilon)-X_2(t)}{\epsilon}\right)\right],$$

which reduces to

$$\lim_{\epsilon\to 0}E\left[X_1(t)\left(\frac{X_2(t+\epsilon)-X_2(t)}{\epsilon}\right)\right],$$

since we know that $\lim_{n\to\infty}E[X[n]] = E[X]$ if $X[n]\to X$ in the m.s. sense. Now, in turn,

$$\begin{aligned}
\lim_{\epsilon\to 0}E\left[X_1(t)\left(\frac{X_2(t+\epsilon)-X_2(t)}{\epsilon}\right)\right] &= \lim_{\epsilon\to 0}\left(\frac{R_{X_1X_2}(t_1,t_2+\epsilon)-R_{X_1X_2}(t_1,t_2)}{\epsilon}\right) \\
&= \frac{\partial}{\partial t_2}R_{X_1X_2}(t_1,t_2)
\end{aligned}$$

2. (a) We need the existence of $\left.\frac{\partial^2 K_{XX}}{\partial t\partial s}\right|_{t=s}$ for the existence of $X'(t)$. For the given covariance function, we get

$$\begin{aligned}
\frac{\partial K_{XX}(t,s)}{\partial s} &= +\sigma^2\frac{\partial}{\partial s}\cos\omega_0(t-s) \\
&= \sigma^2\omega_0\sin\omega_0(t-s).
\end{aligned}$$

Then

$$\frac{\partial}{\partial t}\left(\frac{\partial K_{XX}(t,s)}{\partial s}\right) = \sigma^2\omega_0^2\cos\omega_0(t-s).$$

Thus $\left.\frac{\partial^2 K_{XX}}{\partial t\partial s}\right|_{t=s} = \sigma^2\omega_0^2 < \infty$. Therefore the m.s. derivative exists for all finite t.

(b) We just found that $\frac{\partial^2 K_{XX}}{\partial t\partial s} = \sigma^2\omega_0^2\cos\omega_0(t-s)$ and this is equal to $K_{X'X'}(t,s)$.

(c) $R_{X'X'}(t,s) = K_{X'X'}(t,s)+\mu_{X'}(t)\mu_{X'}^*(s)$. But $\mu_{X'}(t) = \frac{d}{dt}\mu_X(t)$ and $\mu_X(t)$ is constant here. Thus $\mu_{X'}(t) = 0$ and so $R_{X'X'}(t,s) = K_{X'X'}(t,s)$.

3. (a) The condition on $X'(t)$ is the existence of $\left.\frac{d^2R_{XX}(\tau)}{d\tau^2}\right|_{\tau=0}$. The overall condition for $Y = 3X + 2X'$ becomes the existence of $R_{YY}(\tau)|_{\tau=0} = 9R_{XX}(0) - 4R''_{XX}(0)$.

(b)

$$\begin{aligned}
R_{YY}(\tau) &\triangleq E[Y(t+\tau)Y^*(t)] \\
&= E[(3X(t+\tau) + 2X'(t+\tau))(3X(t) + 2X'(t))^*] \\
&= 9R_{XX}(\tau) + 6R_{X'X}(\tau) + 6R_{XX'}(\tau) + 4R_{X'X'}(\tau) \\
&= 9R_{XX}(\tau) + 6R'_{XX}(\tau) + 6R'^*_{XX}(-\tau) - 4R''_{XX}(\tau) \\
&= 9R_{XX}(\tau) - 4R''_{XX}(\tau) \quad \text{since the given } R_{XX} \text{ is real and even,} \\
&= \sigma^2 e^{-(\tau/T)^2}\left[9 + \frac{8}{T^2}\left(1 - \frac{2\tau^2}{T^2}\right)\right].
\end{aligned}$$

This since $R'_{XX}(\tau) = -2\frac{\tau}{T^2}\sigma^2 e^{-(\tau/T)^2}$ and $R''_{XX}(\tau) = \left(\frac{2\tau^2}{T^2}\right)^2 \sigma^2 e^{-(\tau/T)^2} - \frac{2}{T^2}\sigma^2 e^{-(\tau/T)^2}$.

4. The random process $X(t)$ is stationary with mean μ_X and covariance function

$$K_{XX}(\tau) = \frac{\sigma_X^2}{1 + \alpha^2\tau^2}.$$

(a) We have to show that $R_{XX}(\tau)$ has derivatives up to order two. Because $X(t)$ is stationary, the mean is constant, so that $\mu'_X(t) = 0$, therefore

$$\begin{aligned}
\frac{dR_{XX}(\tau)}{d\tau} &= \frac{dK_{XX}(\tau)}{d\tau} \\
&= \frac{-\alpha^2\tau\sigma_X^2}{(1+\alpha^2\tau^2)^2}, \quad \text{which exists for all finite } \tau.
\end{aligned}$$

Next

$$\begin{aligned}
\frac{d^2K_{XX}(\tau)}{d\tau^2} &= \frac{d}{d\tau}\left(\frac{-2\alpha^2\tau}{(1+\alpha^2\tau^2)^2}\right)\sigma_X^2 \\
&= \frac{-2\alpha^2(1+\alpha^2\tau^2)^2 + 2\alpha^2\tau \cdot 2 \cdot 2\alpha^2\tau(1+\alpha^2\tau^2)}{(1+\alpha^2\tau^2)^4}\sigma_X^2 \\
&= \frac{-2\alpha^2(1+\alpha^2\tau^2) + 8\alpha^4\tau^2}{(1+\alpha^2\tau^2)^3}\sigma_X^2 \\
&= \frac{-2\alpha^2(1-3\alpha^2\tau^2)}{(1+\alpha^2\tau^2)^3}\sigma_X^2
\end{aligned}$$

which exists for all τ, and hence, for $\tau = 0$. Therefore the m.s. derivative exists for all finite t.

(b) We have

$$\begin{aligned}
\mu_{\dot{X}}(t) &\triangleq E[\dot{X}(t)] \\
&= E\left[\lim_{\epsilon\to 0}\frac{X(t+\epsilon)-X(t)}{\epsilon}\right] \\
&= \lim_{\epsilon\to 0} E\left[\frac{X(t+\epsilon)-X(t)}{\epsilon}\right], \quad \text{because of m.s. existence,} \\
&= \lim_{\epsilon\to 0}\frac{E[X(t+\epsilon)]-E[X(t)]}{\epsilon} \\
&= \lim_{\epsilon\to 0}\frac{\mu_X-\mu_X}{\epsilon}, \quad \text{by stationarity,} \\
&= \lim_{\epsilon\to 0}\frac{0}{\epsilon} \\
&= \lim_{\epsilon\to 0} 0 \\
&= 0.
\end{aligned}$$

For the covariance of $\dot{X}(t)$, we have

$$\begin{aligned}
K_{\dot{X}\dot{X}}(\tau) &= -\frac{d^2K_{XX}(\tau)}{d\tau^2} \\
&= \frac{2\alpha^2(1-3\alpha^2\tau^2)}{(1+\alpha^2\tau^2)^3}\sigma_X^2, \quad \text{from result in part (a).}
\end{aligned}$$

5. Let $I_1 \triangleq \int_0^T X_1(t)dt$ and $I_2 \triangleq \int_0^T X_2(t)dt$, and consider

$$\begin{aligned}
&\left\|\sum_{i=1}^{N}(a_1X_1(t_i)+a_2X_2(t_i))\Delta t_i - (a_1I_1+a_2I_2)\right\| \\
\leq\ & |a_1|\left\|\sum_{i=1}^{N}X_1(t_i)\Delta t_i - I_1\right\| + |a_2|\left\|\sum_{i=1}^{N}X_2(t_i)\Delta t_i - I_2\right\| \\
\leq\ & |a_1|\delta_1 + \qquad |a_2|\delta_2,
\end{aligned}$$

where both δ_1 and δ_2, both positive, can be taken arbitrarily small by the m.s. existence of the integrals I_1 and I_2. Therefore, it must be that the m.s. limit of $\sum_{i=1}^{N}(a_1X_1(t_i)+a_2X_2(t_i))\Delta t_i$ exists in the limit of $N \nearrow \infty$, ($\Delta t_i \searrow 0$). Also, we see that this m.s. limit must equal $a_1I_1 + a_2I_2$. This m.s. limit is given the *symbol*

$$\int_0^T (a_1X_1(t)+a_2X_2(t))dt$$

6.

7. We have that $I(T) \triangleq \frac{1}{T}\int_0^T X(t)dt, \qquad T > 0.$

(a)

$$\begin{aligned}
E[I(T)] &= E\left[\frac{1}{T}\int_0^T X(t)dt\right] \\
&= \frac{1}{T}E\left[\int_0^T X(t)dt\right] \\
&= \frac{1}{T}\int_0^T E[X(t)]dt \\
&= \frac{1}{T}\int_0^T \mu_X dt \\
&= \mu_X.
\end{aligned}$$

(b)

$$\begin{aligned}
\sigma^2_{I(T)} &= E[(I(T) - E[I(T)])^2] \\
&= E\left[\left(\frac{1}{T}\int_0^T X(t)dt - \mu_X\right)^2\right] \\
& E\left[\left(\frac{1}{T}\int_0^T (X(t) - \mu_X)\, dt\right)^2\right] \\
&= \frac{1}{T^2}\int_0^T\int_0^T K_{XX}(t-s)dtds \\
&= \frac{1}{T^2}\int_{-T}^{T}(T-|\tau|)K_{XX}(\tau)d\tau \\
&= \frac{1}{T}\int_{-T}^{T}\left(\frac{T-|\tau|}{T}\right)K_{XX}(\tau)d\tau,
\end{aligned}$$

where we have made the substitution $\tau \triangleq t-s, \xi \triangleq t+s$, two lines above, and then integrated out in the variable ξ.

8. Given that $X(t)$ is an WSS and Gaussian random process, we have to show that $\dot{Y}(t) = 2X(t)\dot{X}(t)$ in m.s. sense when $Y(t) \triangleq X^2(t)$. We assume that the m.s. derivative $\dot{X}(t)$ exists and that $X(t)$ is second order, i.e. $E[|X^2(t)|] < \infty$. We proceed as follows

$$\begin{aligned}
\dot{Y}(t) &= \lim_{\epsilon\to 0}\left(\frac{Y(t+\epsilon) - Y(t)}{\epsilon}\right) \\
&= \lim_{\epsilon\to 0}\left(\frac{X^2(t+\epsilon) - X^2(t)}{\epsilon}\right) \\
&= \lim_{\epsilon\to 0}\left\{[X(t+\epsilon) + X(t)]\left(\frac{X^2(t+\epsilon) - X^2(t)}{\epsilon}\right)\right\} \\
&\stackrel{?}{=} \lim_{\epsilon\to 0}[X(t+\epsilon) + X(t)] \cdot \lim_{\epsilon\to 0}\left(\frac{X^2(t+\epsilon) - X^2(t)}{\epsilon}\right) \\
&= 2X(t)\dot{X}(t) \quad ?
\end{aligned}$$

To show these last two lines are true in this case, we consider the general case where $X^{(n)}(t) \to X(t)$ and $Y^{(n)}(t) \to Y(t)$ (m.s.). Under what conditions is it then true that $X^{(n)}$(t)Y$^{(n)}$(t)$\to$

$X(t)Y(t)$. We will show that the answer is yes in the Gaussian case. Using the norm notation of the Hilbert space of RVs, we can write for any t

$$\begin{aligned}
\sqrt{E[|X^{(n)}(t)Y^{(n)}(t) - X(t)Y(t)|^2]} &= \left\| X^{(n)}Y^{(n)} - XY \right\| \\
&= \left\| X^{(n)}Y^{(n)} - (X^{(n)}Y - X^{(n)}Y) - XY \right\| \\
&\leq \left\| X^{(n)}Y^{(n)} - X^{(n)}Y \right\| + \left\| X^{(n)}Y - XY \right\|, \text{ by triangle inequality,} \\
&= \left\| X^{(n)}(Y^{(n)} - Y) \right\| + \left\| (X^{(n)} - X)Y \right\|.
\end{aligned}$$

Now,

$$\begin{aligned}
\left\| X^{(n)}(Y^{(n)} - Y) \right\|^2 &= E[|X^{(n)}(t)(Y^{(n)}(t) - Y(t))|^2] \\
&= E[|X^{(n)}(t)|^2|Y^{(n)}(t) - Y(t)|^2] \\
&\leq \sqrt{E[|X^{(n)}(t)|^4} \cdot \sqrt{E[|Y^{(n)}(t) - Y(t)|^4]}, \quad \text{by Schwarz inequality,}
\end{aligned}$$

and similarly for the second term above, $\left\| (X^{(n)} - X)Y \right\|^2 \leq \sqrt{E[|X^{(n)}(t) - X(t)|^4} \cdot \sqrt{E[|Y(t)|^4]}$. In general, we would not know whether these 4^{th} order moments converge or not, but in this Gaussian case, we can use the stated Gaussian 4^{th} order moment property to write

$$E[|X^{(n)}(t) - X(t)|^4] = 3\left(E[|X^{(n)}(t) - X(t)|^2]\right)^2 \to 0 \text{ as } n \to \infty,$$

and similarly $\sqrt{E[|Y^{(n)}(t) - Y(t)|^4} \to 0$ as $n \to \infty$. Then since $E[|Y(t)|^4] < \infty$ and $E[|X^{(n)}(t)|^4] < \infty$, by the same 4^{th} order moment property, we get the general conclusion that $X^{(n)}(\mathrm{t})\mathrm{Y}^{(n)}(\mathrm{t}) \to X(t)Y(t)$ (m.s.) for the Gaussian case.

As for the relevant correlation function relation,

$$\begin{aligned}
R_{YY}(\tau) &= E[X^2(t+\tau)]E[X^2(t)] + 2\left(E[X(t+\tau)X(t)]\right)^2, \\
&\qquad \text{by the same } 4^{\text{th}} \text{order moment property,} \\
&= R_{XX}^2(0) + 2R_{XX}^2(\tau).
\end{aligned}$$

Then

$$\begin{aligned}
R_{\dot{Y}\dot{Y}}(\tau) &= -\frac{d^2}{d\tau^2} R_{YY}(\tau) \\
&= -4\left(R_{XX}(\tau)\frac{d^2 R_{XX}(\tau)}{d\tau^2} + \left(\frac{dR_{XX}(\tau)}{d\tau}\right)^2\right).
\end{aligned}$$

9. Let

$$\begin{aligned}
Y_n'(t) &\triangleq n[Y(t + \frac{1}{n}) - Y(t)] \\
&= n\left[X(t + \frac{1}{n})\cos 2\pi f_0(t + \frac{1}{n}) - X(t)\cos 2\pi f_0 t\right],
\end{aligned}$$

and then use the substitution $\cos 2\pi f_0 t\ \cos 2\pi \frac{f_0}{n} - \sin 2\pi \frac{f_0}{n}\ \sin 2\pi f_0 t$ for the first cosine term, to get

$$\begin{aligned} Y_n'(t) - Y'(t) &= [nX(t+\frac{1}{n}) \cos 2\pi f_0 t\ \cos 2\pi \frac{f_0}{n} - X(t+\frac{1}{n}) \sin 2\pi \frac{f_0}{n}\ \sin 2\pi f_0 t \\ &\quad -X(t)\cos 2\pi f_0 t] + 2\pi f_0 \sin 2\pi f_0 t\ X(t) - \cos 2\pi f_0 t\ X'(t) \\ &= \cos 2\pi f_0 t \left[n \left(X(t+\frac{1}{n}) \cos 2\pi \frac{f_0}{n} - X(t) \right) - X'(t) \right] \\ &\quad + \sin 2\pi f_0 t \left[2\pi f_0\ X(t) - n \sin 2\pi \frac{f_0}{n}\ X(t+\frac{1}{n}) \right]. \end{aligned}$$

So, in the Hilbert space of RVs, for each t,

$$\begin{aligned} \left\| Y_n'(t) - Y'(t) \right\| &\leq |\cos 2\pi f_0 t| \left\| n \left(X(t+\frac{1}{n}) \cos 2\pi \frac{f_0}{n} - X(t) \right) - X'(t) \right\| \\ &\quad + |\sin 2\pi f_0 t| \left\| 2\pi f_0\ X(t) - n \sin 2\pi \frac{f_0}{n}\ X(t+\frac{1}{n}) \right\|, \end{aligned}$$

by use of the triangle inequality and scalar multiplication property $\|aX\| = |a|\,\|X\|$ for X an RV.

It then remains to show that each of the two norms in the above equation, call them (A) and (B), respectively, tends to zero as $n \to \infty$. Since $\sin\theta \approx \theta$ and $\cos\theta \approx 1 - \frac{1}{2}\theta^2$ for θ small, e.g. $2\pi\frac{f_0}{n}$ for n large, we expect that both these norms should tend to zero, since $X'(t)$ is the m.s. derivative of $X(t)$, and hence, $X(t)$ is m.s. continuous. The remainder of this solution is thus devoted to showing this. First we consider (A) and subtract/add $X(t+\frac{1}{n})$ inside the round brackets to obtain

$$\begin{aligned} \left\| n \left(X(t+\frac{1}{n}) \cos 2\pi \frac{f_0}{n} - X(t) \right) - X'(t) \right\| &= \left\| n \left(X(t+\frac{1}{n}) \cos 2\pi \frac{f_0}{n} - X(t+\frac{1}{n}) + X(t+\frac{1}{n}) - X(t) \right) - X'(t) \right\| \\ &\leq \left\| n \left(\cos 2\pi \frac{f_0}{n} - 1 \right) \right\| + \left\| n \left(X(t+\frac{1}{n}) - X(t) \right) - X'(t) \right\| \\ &\leq \left| n (\cos 2\pi \frac{f_0}{n} - 1) \right| \left\| X(t+\frac{1}{n}) \right\| + \left\| n \left(X(t+\frac{1}{n}) - X(t) \right) - X'(t) \right\|. \end{aligned}$$

Now, the first term tends to zero, as $n \to \infty$ since $n(\cos 2\pi\frac{f_0}{n} - 1) = O(\frac{1}{n})$ and the second term is the defining approximant for the m.s. derivative $X'(t)$. Note: We take $\|X(t)\| < \infty$ since $X(t)$ is said to be a second order random process, i.e. $E[|X(t)|^2] < \infty$.

Turning to the second norm (B), we proceed as above, but now subtract/add $2\pi f_0 X(t+\frac{1}{n})$. Then, using the triangle inequality again, we obtain

$$\left\| 2\pi f_0\ X(t) - n \sin 2\pi \frac{f_0}{n}\ X(t+\frac{1}{n}) \right\| \leq |2\pi f_0| \left\| X(t) - X(t+\frac{1}{n}) \right\| + \left| 2\pi f_0 - n \sin 2\pi \frac{f_0}{n} \right| \left\| X(t+\frac{1}{n}) \right\|.$$

Here, the first term on the right-hand side goes to zero as $n \to \infty$ by m.s. continuity, and the second term does via $|2\pi f_0 - n \sin 2\pi\frac{f_0}{n}| \to 0$ using $\sin\theta \approx \theta$ with $\theta = 2\pi\frac{f_0}{n}$ for n large. Thus we have $\|Y_n'(t) - Y'(t)\| \to 0$ for the specified formula for $Y'(t)$ as required to be shown.

10. (a) By definition (10.2-1), since $U(t)$ has independent increments, if we let $t_2 > t_1$,

$$\begin{aligned} E[U(t_1)(U(t_2) - U(t_1))] &= E[U(t_1)]E[U(t_2) - U(t_1)] \\ &= 0, \quad \text{since } U(t) \text{ is zero mean.} \end{aligned}$$

Thus

$$E[U(t_1)U(t_2)] = E[U^2(t_1)].$$

If $t_1 > t_2$, then by the symmetry of the situation, we would get $E[U(t_1)U(t_2)] = E[U^2(t_2)]$. So, in general we have

$$\begin{aligned} E[U(t_1)U(t_2)] &= E[U^2(\min(t_1, t_2))] \\ &= f(\min(t_1, t_2)). \end{aligned}$$

(b) Since $U(t)$ has independent increments, if we have $t_1 < t_2 < t_3 < \cdots < t_n$ and then set

$$\delta \triangleq \min_{1 \le i \le n-1} (t_{i+1} - t_i),$$

then the increments

$$U(t_1 + \delta) - U(t_1), U(t_2 + \delta) - U(t_2), \cdots, U(t_n + \delta) - U(t_n),$$

will be jointly independent for all smaller δ. Then, by dividing by δ, and taking the m.s. limit, we get that the derivatives

$$U'(t_1), U'(t_2), \cdots, U'(t_n),$$

are jointly independent for all $t_1 < t_2 < t_3 < \cdots < t_n$ and all positive integers n. Thus be definition $U'(t)$ is an independent process.

(c) We have already shown:

$$\text{i)} \quad E[U'(t)] = 0, \text{ a constant,}$$

$$\text{ii)} \; R_{U'U'}(t_1, t_2) = 0 = \frac{\partial^2 f(\min(t_1, t_2))}{\partial t_1 \partial t_2}, \quad \text{for } t_1 \neq t_2 .$$

Now

$$\frac{\partial f(\min(t_1, t_2))}{\partial t_1} = \begin{cases} f'(t_1), & t_1 < t_2 , \\ 0, & t_1 > t_2, \end{cases}$$

so

$$\begin{aligned} \frac{\partial}{\partial t_2} \left(\frac{\partial f(\min(t_1, t_2))}{\partial t_1} \right) &= \frac{\partial}{\partial t_2} \left(f'(t_1) u(t_2 - t_1) \right) \\ &= f'(t_1) \delta(t_2 - t_1). \end{aligned}$$

Hence, in order for U' to be WSS, we need either $f'(t)$ to be constant, thus $f(t) = at + b$, i.e. a linear function of time.

11. (a)

$$\int_0^t \int_0^t a(t, \tau_1) R_{XX}(\tau_1, \tau_2) a^*(t, \tau_2) d\tau_1 d\tau_2 \quad < \quad \infty,$$

i.e. exists and is finite.

(b)

$$\mu_Y(t) = \int_0^t a(t, \tau) \mu_X(\tau) d\tau$$

(c)

$$K_{YY}(t, s) = \int_0^t \int_0^s a(t, \tau_1) K_{XX}(\tau_1, \tau_2) a^*(s, \tau_2) d\tau_1 d\tau_2$$

12. (a) Let

$$I = \int_{-\infty}^{t} e^{-(t-s)} X(s) ds,$$

with existence in the m.s. sense if

$$\int_{-\infty}^{t} \int_{-\infty}^{t} e^{-(t-s_1)} e^{-(t-s_2)} R_{XX}(s_1, s_2) ds_1 ds_2 < \infty.$$

(b) Given that $Y(t)$ exists, we have

$$\begin{aligned} R_{YY}(t_1, t_2) &\triangleq E[Y(t_1)Y^*(t_2)] \\ &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} e^{-(t_1-s_1)} e^{-(t_2-s_2)} R_{XX}(s_1, s_2) ds_1 ds_2. \end{aligned}$$

(c) The m.s. existence of the derivative $dY(t)/dt$ depends on whether

$$\frac{\partial^2 R_{XX}(t_1, t_2)}{\partial t_1 \partial t_2} \text{ exists at } t_1 = t_2 = t.$$

We compute this derivative in two steps, as follows:

<u>step1:</u> Take first partial with respect to (wrt) t_1.

$$\begin{aligned} \frac{\partial R_{XX}(t_1, t_2)}{\partial t_1} &= \frac{\partial}{\partial t_1} \left(\int_{-\infty}^{t_1} \int_{-\infty}^{t_2} e^{-(t_1-s_1)} e^{-(t_2-s_2)} R_{XX}(s_1, s_2) ds_1 ds_2 \right) \\ &= e^{-(t_1-t_1)} \int_{-\infty}^{t_2} e^{-(t_2-s_2)} R_{XX}(t_1, s_2) ds_2 \\ &\quad - \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} e^{-(t_1-s_1)} e^{-(t_2-s_2)} R_{XX}(s_1, s_2) ds_1 ds_2. \end{aligned}$$

<u>step2:</u> Take partial of the result in step 2, this time wrt t_2.

$$\begin{aligned} \frac{\partial}{\partial t_2} \left(\frac{R_{XX}(t_1, t_2)}{\partial t_1} \right) &= R_{XX}(t_1, t_2) - \int_{-\infty}^{t_2} e^{-(t_2-s_2)} R_{XX}(t_1, s_2) ds_2 \\ &\quad - \int_{-\infty}^{t_1} e^{-(t_1-s_1)} R_{XX}(s_1, t_2) ds_1 + R_{YY}(t_1, t_2). \end{aligned}$$

So,

$$\begin{aligned} \left. \frac{\partial^2 R_{XX}(t_1, t_2)}{\partial t_1 \partial t_2} \right|_{t_1=t_2=t} &= R_{XX}(t, t) - \int_{-\infty}^{t} e^{-(t-s)} R_{XX}(t, s) ds \\ &\quad - \int_{-\infty}^{t} e^{-(t-s)} R_{XX}(s, t) ds + R_{YY}(t, t) < \infty, \end{aligned}$$

where

$$R_{YY}(t, t) = \int_{-\infty}^{t} \int_{-\infty}^{t} e^{-(t-s_1)} e^{-(t-s_2)} R_{XX}(s_1, s_2) ds_1 ds_2,$$

as in part (a).

13. (a) We have that $X[n] \to Y$ in the m.s. sense, and hence in probability and distribution. Each $X[n]$ is Gaussian distributed with mean μ_n and variance σ_n^2. By m.s. convergence, $\mu_n \to \mu = E[Y]$ and $\sigma_n^2 \to \sigma^2 =Var[Y]$, thus the distribution of $X[n]$ must be converging to $N(\mu, \sigma^2)$. Since each one is $N(\mu_n, \sigma_n^2)$ and since the Gaussian pdf/CDF is continuous in its parameters (for $\sigma^2 > 0$), we get that Y must be distributed as $N(\mu, \sigma^2)$, i.e. be Gaussian too. Note that here, to avoid degeneracy, we need $\sigma^2 > 0$.

(b) Here, we have the vector case $\mathbf{X}[n] \to \mathbf{Y}$ in the m.s. sense. Since m.s. convergence for random vectors implies term-wise m.s. convergence, i.e.

$$E[\|\mathbf{X}[n] - \mathbf{Y}\|^2] = \sum_{i=1}^{K} E[(X_i[n] - Y_i)^2], \qquad \text{(Please don't confuse this vector norm}$$

with the Hilbert space for RVs norm)

we have that

$$\boldsymbol{\mu}_n \triangleq E[\mathbf{X}[n]] \to E[\mathbf{Y}] \triangleq \boldsymbol{\mu}$$

and

$$\mathbf{K}_n \triangleq E[(\mathbf{X}[n] - \boldsymbol{\mu}_n)\,(\mathbf{X}[n] - \boldsymbol{\mu}_n)^T] \to E[(\mathbf{Y} - \boldsymbol{\mu})\,(\mathbf{Y} - \boldsymbol{\mu})^T] \triangleq \mathbf{K}.$$

Then, since the $\mathbf{X}[n]$ are Gaussian distributed as $N(\boldsymbol{\mu}_n, \mathbf{K}_n)$, we have that its m.s. limit random vector $\mathbf{Y}$ is also Gaussian distributed as $N(\boldsymbol{\mu}, \mathbf{K})$.

(c) By definition, a random process is Gaussian if all its finite order distributions are Gaussian. But these are all vectors, thus the result of part (b) applies. So, if $X_n(t) \to Y(t)$ in the m.s. sense, then for any K times t_i, the corresponding K dimensional vectors converge by

$$E[\|\mathbf{X}[n] - \mathbf{Y}\|^2\,] = \sum_{i=1}^{K} E[|X_n(t_i) - Y(t_i)|^2],$$

so the result of part (b) implies that $Y(t)$ is Gaussian distributed at these K samples. Since this result holds for arbitrary positive integers K and for arbitrary times $t_1 < t_2 < \cdots < t_K$, it follows that $Y(t)$ is a Gaussian random process.

14.

15.

16. (a) The *zero-state solution* is

$$\int_{t_0}^{t} e^{-3(t-\tau)} u(\tau) d\tau,$$

where the input $u(\tau)$ is assumed to start at $\tau = t_0$. (This is just $h(t) * u(t)$, where $h(t) = e^{-3t} \times$ (unit step function). The *zero-input solution* is:

$$x_0 e^{-3(t-t_0)}.$$

The *total solution* $x(t)$ is then the sum of these two components:

$$x(t) = x_0 e^{-3(t-t_0)} + \int_{t_0}^{t} e^{-3(t-\tau)} u(\tau) d\tau, \qquad t \geq t_0.$$

(b) Now, for second-order processes, we can also write, for RV X_0 and input random process $U(t)$,

$$X(t) = X_0 e^{-3(t-t_0)} + \int_{t_0}^{t} e^{-3(t-\tau)} U(\tau) d\tau, \qquad t \geq t_0.$$

The first term will be well defined in the m.s. sense if X_0 is a second-order RV, i.e. $E[|X_0|^2] < \infty$. The second term is an m.s. integral, and is well defined if

$$\int_{t_0}^{t} \int_{t_0}^{t} e^{-3(t-\tau_1)} e^{-3(t-\tau_2)} R_{UU}(\tau_1, \tau_2) d\tau_1 d\tau_2 < \infty.$$

Clearly, this will be the case for $R_{UU}(t_1, t_2) = \frac{1}{4} \exp -2|t_1 - t_2|$, which is the correlation function corresponding to the psd $S_{UU}(\omega) = \frac{1}{\omega^2+4}$.

(c) Here are two ways to solve for $R_{UU}(t_1, t_2)$:

(1) Guess the answer is some constant $k \exp -2|\tau|$ and take Fourier transform to get

$$k \left(\int_0^{\infty} e^{(-2-j\omega)\tau} d\tau + \int_{-\infty}^{0} e^{(+2-j\omega)\tau} d\tau \right) = \frac{1}{\omega^2 + 4}.$$

Then solve for $k = \frac{1}{4}$ to fit.

(2) Use Laplace transforms, and write $S_{UU}(s) = \frac{1}{(s+2)(-s+2)}, -2 < \mathrm{Re}(s) < +2$, by substituting $s/j = \omega$. Then do the partial fraction expansion to get

$$\begin{aligned} S_{UU}(s) &= \frac{1}{(s+2)(-s+2)} \\ &= \frac{A}{s+2} + \frac{B}{-s+2}, \end{aligned}$$

with appropriate regions of convergence for each of these terms. We find

$$A = \left. \frac{(s+2)}{(s+2)(-s+2)} \right|_{s=-2} = \frac{1}{4} \quad \text{and} \quad B = \left. \frac{(-s+2)}{(s+2)(-s+2)} \right|_{s=+2} = \frac{1}{4}.$$

Now, for stability, the region of convergence (ROC) of the $\frac{\frac{1}{4}}{s+2}$ term is $\mathrm{Re}(s) > -2$, while the ROC of the $\frac{\frac{1}{4}}{-s+2}$ term is $\mathrm{Re}(s) < +2$. Note that the overlap of these two ROCs is $-2 < \mathrm{Re}(s) < +2$, as required above. From these two terms, we get a causal component $\frac{1}{4} e^{-2\tau} u(\tau)$, where, in this part of the problem, we use u to denote the unit step function, and an anticausal component $\frac{1}{4} e^{+2\tau} u(-\tau)$.

17. (a) We have the following ordinary differential equation for the mean function,

$$\begin{aligned} \dot{\mu}_Y(t) + 2\mu_Y(t) &= \mu_X(t), \quad t > 0, \\ &= 5 \cos 2t, \end{aligned}$$

with $\mu_Y(0) = E[Y(0)] = E[0] = 0$. Using elementary methods, we have the solution to this deterministic linear differential equation as

$$\begin{aligned} \mu_Y(t) &= A e^{-2t} + \mathrm{Re} \left[\frac{5 e^{+j2t}}{j2 + 2} \right] \\ &= 0 \quad \text{at} \quad t = 0. \end{aligned}$$

Now $\frac{5e^{+j2t}}{j2+2} = \frac{5}{2\sqrt{2}}e^{+j(2t-\frac{\pi}{4})}$, so we can write the solution for $t > 0$ simply as

$$\begin{aligned}\mu_Y(t) &= \frac{5}{2\sqrt{2}}\left[\cos(2t-\frac{\pi}{4}) - \frac{1}{\sqrt{2}}e^{-2t}\right] \\ &= \frac{5}{4}\left[\cos 2t + \sin 2t - e^{-2t}\right]. \qquad \text{(alternative form)}\end{aligned}$$

(b) The covariance function is just the correlation function of the centered process $Y_c(t) \triangleq Y(t) - \mu_Y(t)$, which by the linearity of the equation is just the m.s. solution to

$$\frac{dY_c(t)}{dt} + 2Y_c(t) = W(t), \quad t > 0,$$

with initial condition $Y_c(0) = Y(0) - \mu_Y(0) = 0 - 0 = 0$. After solving this first-order differential equation, we can get the covariance function as follows:

$$K_{YY}(t_1,t_2) = \begin{cases} \frac{\sigma^2}{4}e^{-2t_2}(e^{+2t_1} - e^{-2t_1}), & t_1 \le t_2, \\ \frac{\sigma^2}{4}(1-e^{-4t_2})e^{-2(t_1-t_2}, & t_1 > t_2. \end{cases}$$

(c) The variance function $\sigma_Y^2(t) = K_{YY}(t,t) = \frac{\sigma^2}{4}(1-e^{-4t})$, $t > 0$. Using the fact that $Y(t)$ must be Gaussian (since it is the m.s. limit of Gaussian), we have

$$\begin{aligned}P[|Y(t)-\mu_Y(t)| \le 0.1] &= P\left[\left|\frac{Y(t)-\mu_Y(t)}{\sigma_Y(t)}\right| \le \frac{0.1}{\sigma_Y(t)}\right] \\ &= 2\,\mathrm{erf}\left(\frac{0.1}{\sigma_Y(t)}\right) \\ &\ge 0.99, \\ \text{or} \quad & \mathrm{erf}\left(\frac{0.1}{\sigma_Y(t)}\right) \ge 0.495.\end{aligned}$$

Then, from Table, we get approximately

$$\frac{0.1}{\sigma_Y(t)} \ge 2.555, \quad \text{so} \quad \sigma_Y(t) \le \frac{0.1}{2.555}.$$

Then since $\max \sigma_Y^2(t) = \sqrt{\frac{\sigma^2}{4}} = \frac{\sigma}{2}$ for $t > 0$, we need

$$\sigma \le \frac{0.2}{2.555} = 0.078.$$

18.

19.

20. We are given the generalized m.s. differential equation

$$\frac{dX}{dt} + 3X(t) = \frac{dW}{dt} + 2W(t), \quad t \ge t_0,$$

where $X(t_0) = X_0 \perp W(t), \mu_{X_0} = \mu_W(t) = 0$, and psd $S_{WW}(\omega) = 1$ for all ω, i.e. $-\infty < \omega < +\infty$.

(a) So, for $t > t_0$, we get

$$X(t) = Ae^{-3(t-t_0)} + \int_{t_0}^{t} h(t-\nu)W(\nu)d\nu,$$

with $h(t) = \delta(t) - e^{-3t}u(t)$. This becomes

$$X(t) = (X_0 - W(t_0)e^{-3(t-t_0)} + W(t) - \int_{t_0}^{t} e^{-3(t-\nu)}W(\nu)d\nu. \qquad ((\text{A}))$$

(b)

$$\begin{aligned} S_{XX}(\omega) &= |H(j\omega)|^2 S_{WW}(\omega) \\ &= \left|\frac{j\omega+2}{j\omega+3}\right|^2 \cdot 1 \\ &= \frac{\omega^2+4}{\omega^2+9}. \end{aligned}$$

(c)

$$\begin{aligned} S_{XX}(s) &= \frac{(s+2)(-s+2)}{(s+3)(-s+3)} \cdot 1 \\ &= \left(1 - \frac{1}{s+3}\right)\left(1 - \frac{1}{-s+3}\right) \\ &= 1 - \frac{1}{s+3} - \frac{1}{-s+3} + \frac{1}{(s+3)(-s+3)}. \end{aligned}$$

Which implies the autocorrelation function result

$$\begin{aligned} R_{XX}(\tau) &= \begin{cases} \delta(\tau) - e^{-3\tau} + \frac{1}{3+3}e^{-3\tau}, & \tau \geq 0, \\ \delta(\tau) - e^{+3\tau} + \frac{1}{3+3}e^{+3\tau}, & \tau \leq 0, \end{cases} \\ &= \delta(\tau) - \frac{5}{6}e^{-3|\tau|}, \quad -\infty < \tau < +\infty. \end{aligned}$$

(d) By equation (A) above, even if $W(t)$ is Gaussian, still $X(t)$ is not Markov. This because, as we can see from (A), for $t > t_0$, $X(t)$ has direct dependence on more than just $X(t_0)$.

21. We have

$$\widehat{X}(t) = \int_{-\infty}^{+\infty} h(t-\tau)(X(\tau)+N(\tau))d\tau \qquad \text{(if it exists)},$$

so the constraint $E[|\widehat{X}(t)|^2] < \infty$ becomes

(a)

$$E[|\widehat{X}(t)|^2] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(t-\tau_1)h^*(t-\tau_2)[R_{XX}(\tau_1-\tau_2) + R_{NN}(\tau_1-\tau_2)]d\tau_1 d\tau_2 < \infty,$$

where we have used the orthogonality of the X and N processes.

(b) In the Fourier transform domain we have

$$\begin{aligned}
E[|\widehat{X}(t)|^2] &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_{\widehat{X}\widehat{X}}(\omega)d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} |H(\omega)|^2 \left(S_{XX}(\omega)+S_{NN}(\omega)\right) d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left(\frac{S_{XX}(\omega)}{S_{XX}(\omega)+S_{NN}(\omega)}\right)^2 \left(S_{XX}(\omega)+S_{NN}(\omega)\right) d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{S_{XX}^2(\omega)}{S_{XX}(\omega)+S_{NN}(\omega)} d\omega < \infty.
\end{aligned}$$

(c) Since the WSS process $X(t)$ is second-order, i.e. $E[|X(t)|^2] < \infty$, we know $\frac{1}{2\pi}\int_{-\infty}^{+\infty} S_{XX}(\omega)d\omega < \infty$. Thus, we can show the condition in part (b) as follows

$$\begin{aligned}
E[|\widehat{X}(t)|^2] &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_{XX}(\omega)\left(\frac{S_{XX}(\omega)}{S_{XX}(\omega)+S_{NN}(\omega)}\right) d\omega \\
&\le \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_{XX}(\omega)d\omega, \quad \text{since } 0 \le \frac{S_{XX}(\omega)}{S_{XX}(\omega)+S_{NN}(\omega)} \le 1, \\
&< \infty, \quad \text{since } X \text{ is second-order.}
\end{aligned}$$

22. We have two hypotheses

$$\left.\begin{array}{ll} H_0: & R(t) = W(t) \\ H_1: & R(t) = A + W(t) \end{array}\right\} 0 \le t \le T$$

$$\Lambda = \int_0^T R(t)dt$$

(a) (i) Under hypothesis H_0 :

$$\begin{aligned}
E[\Lambda|H_0] &= E\left[\int_0^T R(t)dt|H_0\right] \\
&= E\left[\int_0^T W(t)dt\right] \\
&= E[0] = 0.
\end{aligned}$$

(ii) Under hypothesis H_1 :

$$\begin{aligned}
E[\Lambda|H_1] &= E\left[\int_0^T R(t)dt|H_1\right] \\
&= E\left[\int_0^T \left(A+W(t)\right)dt\right] \\
& \quad E\left[\int_0^T Adt\right] \\
&= E[AT] = AT.
\end{aligned}$$

(b) For the variances, under H_0 :

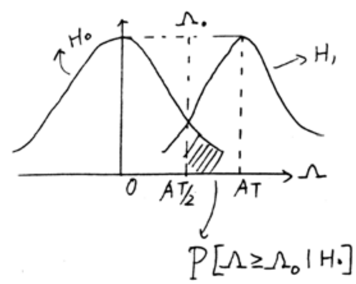
$$\begin{aligned}
\sigma_\Lambda^2 &= E[(\Lambda-\mu_\Lambda)^2|H_0] \\
&= E\left[\left(\int_0^T W(t)dt - \int_0^T \mu_W(t)dt\right)^2 |H_0\right] \\
&= E\left[\left(\int_0^T W(t)dt - 0\right)^2\right] \\
&= E\left[\int_0^T\int_0^T W(t_1)W(t_2)dt_1dt_2\right] \\
&= \int_0^T\int_0^T \sigma^2\delta(t_1-t_2)dt_1dt_2 \\
&= \sigma^2 T.
\end{aligned}$$

Under hypothesis H_1, the variance is the same since there is only a shift in the DC value. Thus under each hypothesis $\sigma_\Lambda^2 = \sigma^2 T$.

(c)

$$P[\Lambda \geq \Lambda_0|H_0] = \frac{1}{\sqrt{2\pi\sigma_\Lambda^2}}\int_{\Lambda_0}^\infty e^{-\frac{\alpha^2}{2\sigma_\Lambda^2}}d\alpha,$$

where $\Lambda_0 \triangleq AT/2$. Let $\frac{\alpha}{\sigma_\Lambda} = \eta$, then $d\alpha = \sigma_\Lambda d\eta$. Also $\alpha = \frac{AT}{2}$, which implies $\eta = \frac{AT}{2\sigma_\Lambda}$.



Then

$$\begin{aligned}
P[\Lambda \geq \Lambda_0|H_0] &= \frac{\sigma_\Lambda}{\sqrt{2\pi}\sigma_\Lambda}\int_{\frac{AT}{2\sigma_\Lambda}}^\infty e^{-\frac{\eta^2}{2}}d\eta \\
&= \frac{1}{2} - \operatorname{erf}\left(\frac{AT}{2\sigma_\Lambda}\right) \\
&= \frac{1}{2} - \operatorname{erf}\left(\frac{AT}{2\sigma\sqrt{T}}\right) \\
&= \frac{1}{2} - \operatorname{erf}\left(\frac{A\sqrt{T}}{2\sigma}\right).
\end{aligned}$$

23. (a) Let

$$I_T \triangleq \frac{1}{2T}\int_{-T}^{+T} X(t)dt.$$

Then, if $E[I_T] = E[X(t)] = \mu_X$, a constant, and if

$$\sigma_{I_T}^2 \triangleq E[(I_T - \mu_X)^2] \underset{T\to\infty}{\longrightarrow} 0,$$

then $X(t)$ is ergodic in the mean.

(b) Using the sufficient condition, i.e.

$$\begin{aligned}\int_{-\infty}^{+\infty} K_{XX}(\tau)d\tau &< \infty, \text{ or}\\ \int_{-\infty}^{+\infty} R_{XX}(\tau)d\tau &< \infty,\end{aligned}$$

in the case of zero mean, we have

$$\begin{aligned}\int_{-\infty}^{+\infty} \sigma^2 e^{-\alpha|\tau|}\cos(2\pi f\tau)d\tau &= 2\int_0^{+\infty} \sigma^2 e^{-\alpha\tau}\left(\frac{e^{j2\pi f\tau}+e^{-j2\pi f\tau}}{2}\right)d\tau\\ &= \sigma^2\left[\frac{e^{(-\alpha+j2\pi f)\tau}}{-\alpha+j2\pi f} - \frac{e^{-(\alpha+j2\pi f)\tau}}{\alpha+j2\pi f}\right]\Bigg|_0^{\infty}\\ &= -\sigma^2\left[\frac{1}{-\alpha+j2\pi f} - \frac{1}{\alpha+j2\pi f}\right]\\ &= \frac{2\alpha\sigma^2}{\alpha^2+(2\pi f)^2} < \infty.\end{aligned}$$

Therefore, the given random process $X(t)$ is ergodic in the mean.

24. Since $X(t)$ is ergodic in the mean,

$$\frac{1}{2T}\int_{-2T}^{+2T}\left(1-\frac{|\tau|}{2T}\right)K_{XX}(\tau)d\tau \underset{T\to\infty}{\longrightarrow} 0.$$

Now, as shown in Theorem 10.4-4, this can only happen together with $K_{XX}(\tau) \underset{|\tau|\to\infty}{\longrightarrow}$ a constant. If this constant is zero, i.e.

$$\lim_{|\tau|\to\infty} K_{XX}(\tau) = 0.$$

But, always we have

$$R_{XX}(\tau) = K_{XX}(\tau) + |\mu_X(\tau)|^2,$$

so

$$\lim_{|\tau|\to\infty} R_{XX}(\tau) = |\mu_X(\tau)|^2$$

under the stated conditions.

25. (a) Let $X_c = X - \mu_X$, then

$$\begin{aligned}E[|\widehat{M} - \mu_X|^2] &= \frac{1}{N^2}E\left[\left|\sum_{n=1}^{N} X_c[n]\right|^2\right]\\ &= \frac{1}{N^2}\sum_{n,m=1}^{N} K_{XX}[n-m] \underset{N\to\infty}{\longrightarrow} 0.\end{aligned}$$

(b)

$$\begin{aligned}
E[|\widehat{M}-\mu_X|^2] &= \frac{1}{N^2}[NK_{XX}[0]+(N-1)(K_{XX}[+1]+K_{XX}[-1])+ \\
&\quad +(N-2)(K_{XX}[+2]+K_{XX}[-2])+\cdots] \\
&= \frac{1}{N}\sum_{m=-N}^{+N}\left(1-\frac{|m|}{N}\right)K_{XX}[m] \underset{N\to\infty}{\to} 0.
\end{aligned}$$

(c) We calculate

$$\begin{aligned}
\sum_{m=-\infty}^{+\infty}|K_{XX}[m]| &\leq 5\sum_{m=-\infty}^{+\infty}0.9^{|m|}+15\sum_{m=-\infty}^{+\infty}0.8^{|m|} \\
&< \infty.
\end{aligned}$$

So,

$$\frac{1}{N}\sum_{m=1}^{N}|K_{XX}[m]| \underset{N\to\infty}{\to} 0,$$

which implies that $X[n]$ is ergodic in the mean.

26. Define the one-parameter covariance function $K_{XX}(\tau) \triangleq K_{XX}(s+\tau,s)=\sigma^2\cos\omega_0\tau$, which is seen to be periodic in τ and independent of s. Then since the mean μ_X is constant, we have a WSS periodic random process. For these processes, the Fourier series expansion coefficients are orthogonal, i.e. the Fourier series basis set is also the Karhunen-Loeve basis set. The period of $K_{XX}(\tau)$ is $T \triangleq \frac{2\pi}{\omega_0}$. Hence, any interval of the time axis of width T will do for the expansion.

27. By the equation (10.5-3),

$$\int_{-T/2}^{T/2}K_{XX}(t_1,t_2)\phi_1(t_2)\,dt_2=\lambda_1\phi_1(t_1),$$

we have

$$\begin{aligned}
\int_{-T/2}^{T/2}\phi_1(t)\phi_2^*(t)dt &= \frac{1}{\lambda_1}\int_{-T/2}^{T/2}\int_{-T/2}^{T/2}K_{XX}(t,t_2)\phi_1(t_2)\phi_2^*(t)dtdt_2 \\
&= \frac{1}{\lambda_1}\int_{-T/2}^{T/2}\phi_1(t_2)\left(\int_{-T/2}^{T/2}K_{XX}^*(t_2,t)\phi_2^*(t)dt\right)dt_2 \\
&= \frac{1}{\lambda_1}\int_{-T/2}^{T/2}\phi_1(t_2)\lambda_2^*\phi_2^*(t_2)dt_2 \\
&= \frac{\lambda_2^*}{\lambda_1}\int_{-T/2}^{T/2}\phi_1(t)\phi_2^*(t)dt \\
&= \frac{\lambda_2}{\lambda_1}\int_{-T/2}^{T/2}\phi_1(t)\phi_2^*(t)dt, \text{ since the } \lambda_i \text{ are real.}
\end{aligned}$$

Because $\lambda_1 \neq \lambda_2$, it must be that the indicated integral is zero.

28.

29.

30.

31.

32.

33.

34.

35.

36.

37.

38.

39.

40.

41.

42.

43. (a) $y_1 = h * x_1 + g * x_2$, so

$$\begin{aligned} S_{Y_1X_1}(\omega) &= FT\{R_{Y_1X_1}(\tau)\} \\ &= FT\{E[(h * X_1)(t+\tau)X_1^*(t)] + E[(g * X_2)(t+\tau)X_1^*(t)]\} \\ &= FT\{h(\tau) * R_{X_1X_1}(\tau) + g(\tau) * R_{X_2X_1}(\tau)\}, \\ &= FT\{h(\tau) * R_{X_1X_1}(\tau)\}, \quad \text{since } R_{X_2X_1}(\tau) = 0 \text{ since } X_2 \perp X_1, \\ &= H(\omega)\, S_{X_1X_1}(\omega). \end{aligned}$$

(b)

$$\begin{aligned} S_{Y_2X_2}(\omega) &= FT\{E[Y_2(t+\tau)X_2^*(t)]\} \\ &= FT\{b(\tau) * R_{X_2X_2}(\tau)\}, \quad \text{since } \; U \perp X_2 \\ &= B(\omega) S_{X_2X_2}(\omega). \end{aligned}$$

(c)

$$\begin{aligned} E[Y_1(t+\tau)Y_2^*(t)] &= E[(h * X_1 + g * X_2)(t+\tau)(U^* + b^* * X_2^*)(t)] \\ &= E[(g * X_2)(t+\tau)(b^* * X_2^*)(t)], \quad \text{since } \; U \perp X_1 \text{ and } X_2, \\ &= g(\tau) * b^*(-\tau) * R_{X_2X_2}(\tau). \end{aligned}$$

Thus $S_{Y_1Y_2}(\omega) = G(\omega) B^*(\omega) S_{X_2X_2}(\omega)$.

44. (a) Since the noise process is Gaussian, the K-L expansion ensures independence of the transformed coefficients. The other $R_k's$ are thus independent of R_{k_o}, which is the only one containing the message. Thus

$$P[R_{k_o} \leq r | \{ \text{ all other } R_k's \}] = P[R_{k_o} \leq r].$$

(b) Since λ_k is the noise mean-square level on basis function (channel) k, we want the smallest λ_k for the signaling channel. So we want $k_o = \infty$. Of course, practical conditions would intercede in reality, forcing a lower finite choice.

45. (a) We need the existence of $\left.\frac{\partial^2 K_{XX}}{\partial t \partial s}\right|_{t=s}$ for the existence of $X'(t)$. Calculating, we find

$$\begin{aligned} \frac{\partial K_{XX}(t,s)}{\partial s} &= \sigma^2 \frac{\partial}{\partial s} \cos \omega_0 (t-s) \\ &= \sigma^2 \omega_0 \sin \omega_0 (t-s). \end{aligned}$$

Then

$$\begin{aligned} \left.\frac{\partial}{\partial t} \left(\frac{\partial K_{XX}(t,s)}{\partial s} \right)\right|_{t=s} &= \sigma^2 \omega_0^2 \cos \omega_0 (t-s)\big|_{t=s} \\ &= \sigma^2 \omega_0^2 < \infty. \end{aligned}$$

Since this value is finite, the m.s. derivative exists for all t.

(b) From part a), we have

$$\begin{aligned} \frac{\partial^2 K_{XX}(t,s)}{\partial t \partial s} &= \sigma^2 \omega_0^2 \cos \omega_0 (t-s) \\ &= K_{X'X'}(t,s). \end{aligned}$$

46. (a)

$$\begin{aligned} E[|X(t+\epsilon) - X(t)|^2] &= R_{XX}(t+\epsilon, t+\epsilon) - R_{XX}(t+\epsilon, t) \\ &\quad - R_{XX}(t, t+\epsilon) + R_{XX}(t,t) \\ &\longrightarrow 0 \quad \text{as} \quad \epsilon \searrow 0, \end{aligned}$$

for $R_{XX}(t,s)$ continuous in t and s.

(b) The integral $I = \int_a^b X(t)dt$ exists in the m.s. sense, for $-\infty < a \leq b < +\infty$, if the two-dimensional deterministic integral

$$\int_a^b \int_a^b R_{XX}(t,s) dt ds \text{ exists.}$$

But since $R_{XX}(t,s)$ is here assumed to be continuous, by the quoted fact, this ordinary 2-D Riemann integral will exist for all finite values of $a \leq b$.

47.

48.

Chapter 11 Solutions

1. For minimum variance, we want to minimize diagonal terms of

$$\epsilon^2 \triangleq \overline{(\hat{Y}-Y)(\hat{Y}-Y)^T},$$

where $\hat{Y} = AX$.

$$\epsilon^2 = \overline{(AX-Y)(AX-Y)^T} = AK_1A^T + K_2 - AK_{12} - K_{21}A^T.$$

Now write $A = A_0 + \delta$; is $\delta = 0$ for minimum variance?

$$\begin{aligned}
\epsilon^2 &= (A_0+\delta)K_1(A_0+\delta)^T + K_2 - (A_0+\delta)K_{12} - K_{21}(A_0+\delta)^T \\
&= A_0K_1A_0^T + K_2 - A_0K_{12} - K_{21}A_0^T + \delta K_1\delta^T + A_0K_1\delta^T + \delta K_1A_0^T - \delta K_{12} - K_{21}\delta^T \\
&= \overline{(A_0X-Y)(A_0X-Y)^T} + \delta K_1\delta^T + A_0K_1\delta^T + \delta K_1A_0^T - \delta K_{12} - K_{21}\delta^T \\
&= \overline{(A_0X-Y)(A_0X-Y)^T} + \delta K_1\delta^T + K_{21}\delta^T + \delta K_{12} - \delta K_{12} - K_{21}\delta^T \\
&= \overline{(A_0X-Y)(A_0X-Y)^T} + \delta K_1\delta^T + 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\text{Trace}(\epsilon^2) &= \text{Trace}\left[(Y-A_0X)(Y-A_0X)^T\right] + \text{Trace}\left(\delta K_1\delta^T\right) \\
&= \sum_{i=1}^{n}\left[y_i - \sum_j a_{ij}^{(0)}x_j\right]^2 + \text{Trace}\left[\delta CC^T\delta^T\right] \\
&\geq 0.
\end{aligned}$$

$\text{Trace}\left[\delta CC^T\delta^T\right] = \text{Trace}\left[\delta C(\delta C)^T\right] \geq 0$ where we use factorization $K_1 = CC^T$. Clearly set $\delta = 0$ for $\text{Trace}(\epsilon^2)$ to be minimum.

2. Suppose $Y_1 = Y - \mu_2$ and $X_1 = X - \mu_1$. Then we know that $E[Y_1] = 0 = E[X_1]$, $K_1 = E[X_1X_1^T], K_2 = E[Y_1Y_1^T], K_{12} = E[X_1Y_1^T], K_{21} = E[Y_1X_1^T]$. Hence from problem 11.1, the minimum-variance estimator Y_1 of the form $\hat{Y}_1 = AX_1$ is given by

$$\hat{Y}_1 = K_{21}K_1^{-1}(X-\mu_1).$$

But is $\hat{Y}_1$ equal to $\hat{Y} - \mu_2$?

Write $Y = Y_1 + \mu_2$. Let β be such that $\hat{Y} = \hat{Y}_1 + \beta$. Then

$$\begin{aligned}
\epsilon_T^2 &= \overline{(\hat{Y}-Y)(\hat{Y}-Y)^T} \\
&= \overline{[(\hat{Y}_1-Y_1)+(\beta-\mu_2)][(\hat{Y}_1-Y_1)+(\beta-\mu_2)]^T} \\
&= \overline{(\hat{Y}_1-Y_1)(\hat{Y}_1-Y_1)^T} + (\beta-\mu_2)(\beta-\mu_2)^T \\
&\quad (\text{because } \overline{(\hat{Y}_1-Y_1)} = E[K_{21}K_1^{-1}X_1] - E[Y_1] = 0-0 = 0) \\
&= \epsilon^2 + (\beta-\mu_2)(\beta-\mu_2)^T
\end{aligned}$$

$\text{tr}\epsilon_T^2 = \text{tr}\epsilon^2 + \text{tr}[(\beta-\mu_2)(\beta-\mu_2)^T] \geq 0$ is minimum when $\beta = \mu_2$. Hence

$$\hat{Y} = \hat{Y}_1 + \mu_2 = \mu_2 + K_{21}K_1^{-1}(X-\mu_1).$$

3. Given $\epsilon^2 = E[(X - E[X|Y])^2] = E[X(X - E[X|Y])] - E[E[X|Y][X - E[X|Y]]]$.
However, since the estimate is orthogonal to the error, we have

$$E[E[X|Y][X - E[X|Y]]] = 0. \tag{1}$$

Therefore,

$$\epsilon^2 = E[X[X - E[X|Y]]]. \tag{2}$$

By the same reasoning, from Eq. 1: $E[E[X|Y][X - E[X|Y]]] = 0 \implies E[XE[X|Y]] = E[[E[X|Y]]^2]$. But from Eq. 2,

$$\epsilon^2 = E[X^2] - E[XE[X|Y]] = E[X^2] - E[(E[X|Y])^2].$$

Now let us extend this to random N-vectors.

$$\begin{aligned} \epsilon^2 &= E[[X - E[X|Y]][X - E[X|Y]]^T] & (3) \\ &= E[XX^T] - E[XE[X|Y]^T] - E[E[X|Y]X^T] + E[E[X|Y]E[X|Y]^T]. & (4) \end{aligned}$$

However, using orthogonality again,

$$E[E[X|Y][X - E[X|Y]^T]] = 0$$

or

$$E[E[X|Y]X^T] = E[E[X|Y]E[X|Y]^T]. \tag{5}$$

Using Eq. 3 and Eq. 5 together,

$$\epsilon^2 = E[XX^T] - E[XE[X|Y]^T] = E[X[X - E[X|Y]]^T]. \tag{6}$$

Also:

$$E[XE[X|Y]^T] = E[E[X|Y]E^T[X|Y]]. \tag{7}$$

Using Eq. 6 and 7 together, we have

$$\epsilon^2 = E[XX^T] - E[E[X|Y]E^T[X|Y]].$$

4. $G[N] \triangleq E[X[n]|Y_N]$, $Y_N \triangleq [Y[n], \ldots, Y[n-N]]$.
From problem 8.55, we know $G[N]$ is a Martingale sequence in the parameter N.

$$\begin{aligned} E[X^2[n]] &= E[(X[n] - E[X[n]|Y_N] + E[X[n]|Y_N])^2] \\ &= E[(X[n] - E[X[n]|Y_N])^2] + E[E^2[X[n]|Y_N]]. \\ &\quad \text{(by the orthogonality principle)} \end{aligned}$$

$E[E^2[X[n]|Y_N]] \leq E[X^2[n]]$

$$\sigma_G^2[N] = E[E^2[X[n]|Y_N]] - E^2[E[X[n]|Y_N]] \leq E[X^2[n]] - E^2[X[n]].$$

Therefore, $\sigma_G^2[N] \leq \sigma_X^2[n]$. Let $C \triangleq \sigma_X^2[n] \implies \sigma_G^2[N] \leq C$ for all N. By the Martingale convergence theorem 8.8-4, we can conclude that the limit $\lim_{N\to\infty} E[X[n]|Y_N]$ exists with probability 1.

5. The modified Theorem 11.1-3 is given as:
Modified theorem: The LMMSE estimate of the zero-mean sequence $X[n]$ based on the zero-mean random sequence $Y[n]$'s $(p+1)$ most recent terms is

$$\hat{E}\left\{X[n]|Y[n],\ldots,Y[n-p]\right\} = \sum_{i=0}^{p} a_i^{(p)} Y[n-i],$$

where the $a_i^{(p)}$ satisfy the orthogonality condition

$$\left[X[n] - \sum_{i=0}^{p} a_i^{(p)} Y[n-i]\right] \perp Y[n-k], 0 \leq k \leq p.$$

Further, the LMMSE is given as

$$\epsilon_{\min}^{2(p)} = E\left\{|X[n]|^2\right\} - \sum_{i=0}^{p} a_i^{(p)} E\left\{Y[n-i]X^*[n]\right\}.$$

(a) Eq. (11.1-25) changes to

$$E[X[n]Y^*[n-k]] = \sum_{i=0}^{p} a_i^{(p)} E\left\{Y[n-i]Y^*[n-k]\right\}, 0 \leq k \leq p$$

via the data modification of $\{Y[n],\ldots,Y[0]\}$ being replaced by $\{Y[n],\ldots,Y[p]\}$. The equation (11.2-4) changes via

$$a^{(p)} \triangleq \left[a_0^{(p)},\ldots,a_p^{(p)}\right]$$

and

$$\underline{Y} \triangleq \left[Y[n], Y[n-1],\ldots,Y[n-p]\right],$$

thereby reversing time from the previous $\underline{Y}$. The cross-covariance vector $\underline{K}_{X\underline{Y}}$ then becomes

$$\underline{K}_{X\underline{Y}} = E\left\{X[n]\underline{Y}^*[n],\ldots,X[n]Y^*[n-p]\right\}$$

and the matrix vector equation (9.1-26) remains essentially the same at

$$\underline{a}^{(p)T} = \underline{K}_{X\underline{Y}}\underline{K}_{\underline{YY}}^{-1}.$$

The entries in $\underline{K}_{\underline{YY}}^{-1}$ are given as

$$\left(\underline{K}_{\underline{YY}}\right)_{ij} = E\left[Y[n-i+1]Y^*[n-j+1]\right], 1 \leq i,j \leq p+1.$$

(b) The modified Eq. (11.1-27) is the same as before with the new $\underline{K}_{X\underline{Y}}$ and $\underline{K}_{\underline{YY}}$ inserted into the old version. Nothing else changes

$$\epsilon_{\min}^{2(p)} = \sigma_X^2(n) - \underline{K}_{X\underline{Y}}\underline{K}_{\underline{YY}}^{-1}\underline{K}_{X\underline{Y}}^T.$$

6. $X[n]$ is defined as

$$X[n] \triangleq -\sum_{k=1}^{n} \binom{k+2}{2} X[n-k] + W[n],$$

for $n = 1, 2, \ldots$ with $X[0] = W[0], W[n]$ is Gaussian noise with zero mean and unit variance.

(a) $W[n]$ is the innovation sequence for $X[n]$ because

- $W[n]$ is a white (uncorrelated) sequence.
- It is defined as a causal invertible linear transformation on $X[n]$.

(b) A simple substitution will show the result. From the definition of $X[n]$,

$$W[n] = X[n] + \sum_{k=1}^{n} \binom{k+2}{2} X[n-k] = \sum_{k=0}^{n} \binom{k+2}{2} X[n-k].$$

Then, $W[n] - 2W[n-1] + 3W[n-2] - W[n-3]$
$= \sum_{k=0}^{n} \binom{k+2}{2} X[n-k] - 3\sum_{k=0}^{n-1} \binom{k+2}{2} X[n-1-k] + 3\sum_{k=0}^{n-2} \binom{k+2}{2} X[n-2-k] - \sum_{k=0}^{n-3} \binom{k+2}{2} X[n-3-k]$
$= X[n]$.

(c) $X[n]$ is Gaussian since $W[n]$ is Gaussian. $\hat{X}[12|10]$
$\triangleq \hat{E}[X[12]|X[0], X[1], \ldots, X[10]]$
$= \hat{E}[X[12]|W[0], W[1], \ldots, W[10]]$
$= \hat{E}[W[12] - 3W[11] + 3W[10] - W[9]|W[0], W[1], \ldots, W[10]]$
$= \hat{E}[W[12]|W[0], W[1], \ldots, W[10]] - 3\hat{E}[W[11]|W[0], W[1], \ldots, W[10]]$
$+ 3\hat{E}[W[10]|W[0], W[1], \ldots, W[10]] - \hat{E}[W[9]|W[0], W[1], \ldots, W[10]]$
$= 3W[10] - W[9]$
$=$ (because $W[12] \perp [W[0], W[1], \ldots, W[10]]$ and $W[11] \perp [W[0], W[1], \ldots, W[10]]$)
For M.S. prediction error,

$$\begin{aligned}
E[(X[12] - \hat{X}[12|10])^2] &= E[((W[12] - 3W[11] + 3W[10] - W[9]) - (3W[10] - W[9]))] \\
&= E[(W[12] - 3W[11])^2] \\
&= E[W^2[12]] + 9E[W^2[11]] - 3E[W[12]W[11]] \\
&= 1 + 9 + 0 = 10.
\end{aligned}$$

$E[W[12]W[11]] = 0$ because $W[12] \perp W[11]$ and $W[n]$ is zero-mean.

7. (a) The innovations sequence is clearly $W[n]$, because $X[n] = X[n-1] + W[n]$.

(b)

$$\hat{X}[n|n] = \hat{X}[n-1|n-1] = G_n \left(Y[n] - \hat{X}[n-1|n-1] \right).$$

(c)

$$G_n = \epsilon^2[n] \left[\epsilon^2[n] + \sigma_v^2 \right]^{-1}; n \geq 1.$$
$$\epsilon^2[n] = \epsilon^2[n-1](1 - G_{n-1}) + 1; n \geq 1.$$

$\epsilon^2[0] = E[X^2[0]] = 0.$

8. Given $2Y[n+2] + Y[n+1] + Y[n] = 2W[n]$, $W[n] \sim N(0,1)$, $Y[0] = 0$, $Y[1] = 1$, we have for state-space representation

$$\begin{bmatrix} Y[n+2] \\ Y[n+1] \end{bmatrix} = \begin{bmatrix} -0.5 & -0.5 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y[n+1] \\ Y[n] \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} W[n].$$

If the state vector $X[n]$ is defined as

$$\underline{X}[n] \triangleq \begin{bmatrix} Y[n+2] \\ Y[n+1] \end{bmatrix},$$

we have

$$\underline{X}[n] = \begin{bmatrix} -0.5 & -0.5 \\ 1 & 0 \end{bmatrix} \underline{X}[n-1] + \begin{bmatrix} 1 \\ 0 \end{bmatrix} W[n].$$

Also, $E[2Y[n+2] + Y[n+1] + Y[n]] = 2E[W[n]]$, and so we have

$$2\mu_Y[n+2] + \mu_Y[n+1] + \mu_Y[n] = 2\mu_W[n]$$

. Obviously, $\mu_Y[0] = 0$ and $\mu_Y[1] = 1$. Then,
$2\mu_Y[2] = -\mu_Y[1] - \mu_Y[0] = -1$ or $\mu_Y[2] = -0.5$.
$2\mu_Y[3] = -\mu_Y[2] - \mu_Y[1] = -0.5$ or $\mu_Y[3] = -0.25$. etc.
We also have $E\left[[2Y[n+2] + Y[n+1] + Y[n]]Y[n]\right] = 2E[W[n]Y[n]] = 0$.
$E\left[[2Y[n+2] + Y[n+1] + Y[n]]Y[n+1]\right] = 2E[W[n]Y[n+1]] = 0$.
Therefore, we can calculate $R[n_1, n_2]$ for all n_1, n_2. Note for initial conditions, we have

(a) $R_Y[0,0] = 0 = R_Y[0,1] = R_Y[1,0]$, and $R_Y[1,1] = 1, R_Y[2,1] = -0.5$ etc.

(b) $R_Y[n_1, n_2] = R_Y[n_2, n_1]$ since Y is a real sequence.

9. (a) From the equation, we have $X[n] = AX[n-1] + BW[n]$. We have

$$\mu_X[n] \triangleq E[X[n]] = E\left[AX[n-1] + BW[n]\right] = A\mu_X[n-1] + B\mu_W[n].$$

From the equation, $Y[n] = X[n] + V[n]$. We also have

$$\mu_Y[n] \triangleq E[Y[n]] = E[X[n]] + E[V[n]] = \mu_X[n] + \mu_V[n].$$

(b) Since $\mu_X[n], \mu_Y[n]$ are deterministic variables, we have

$$\begin{aligned} \hat{X}[n|n] &\triangleq E[X[n]|Y[0], \ldots, Y[n]] \\ &= E[X_C[n] + \mu_X[n]|Y[0], \ldots, Y[n]] \\ &= E[X_C[n]|Y[0], \ldots, Y[n]] + \mu_X[n] \\ &= E[X_C[n]|Y_C[0] + \mu_Y[0], \ldots, Y_C[n] + \mu_Y[n]] + \mu_X[n] \\ &= E[X_C[n]|Y_C[0], \ldots, Y_C[n]] + \mu_X[n] \\ &= \hat{X}_C[n|n] + \mu_X[n]. \end{aligned}$$

(c) Since $X_C[n] = AX_C[n-1] + BW_C[n]$ and $Y_C[n] = X_C[n] + V_C[n]$ are the same as in (11.2-6 and 7), we can use the same estimate equation to estimate $\hat{X}_C[n|n]$, i.e.,

$$\hat{X}_C[n|n] = A\hat{X}_C[n-1|n-1] + G_n\left[Y_C[n] - A\hat{X}[n-1|n-1]\right].$$

So,

$$\hat{X}_C[n|n] = A\hat{X}[n-1|n-1] + G_n\left[Y[n] - A\hat{X}[n-1|n-1]\right] - A\mu_X[n-1] - G_n\mu_Y[n] + G_nA\mu_X[n-1].$$

Therefore,

$$\hat{X}_C[n|n] = A\hat{X}[n-1|n-1] + G_n\left[Y[n] - A\hat{X}[n-1|n-1]\right] - (G_nA - A)\mu_X[n-1] - G_n\mu_Y[n].$$

(d) Since the gain and error covariance equations just depend on $\sigma_V^2[n], \sigma_W^2[n]$ and dynamical model's coefficients, the gain and error covariance equations do not change.

10. $Y[n] = C_n X[n] + V[n]$

We arrive at following equation simply by following the procedure developed in the book.

$$\hat{X}[n] = A_n \left[(I - G_{n-1}C_{n-1})\hat{X}[n-1] + G_{n-1}Y[n-1] \right],$$

with

$$G_n = E\left[X[n]\tilde{Y}^T[n]\right]\left[\sigma_{\tilde{Y}}^2[n]\right]^{-1}$$

and $\tilde{Y}[n] = Y[n] - C_n\hat{X}[n]$ where $\hat{X}[n] \triangleq \hat{X}[n|n-1]$. However,

$$\begin{aligned} E[X[n]\tilde{Y}^T[n]] &= E\left[X[n](Y[n] - C_n\hat{X}[n])^T\right] \\ &= E\left[X[n](C_nX[n] + V[n] - C_n\hat{X}[n])^T\right] \\ &= E\left[X[n](X[n] - \hat{X}[n])^T\right] C_n^T. \end{aligned}$$

By the orthogonality principle, $E\left[\hat{X}[n](X[n] - \hat{X}[n])^T\right] = 0$. So

$$E[X[n]\tilde{Y}^T[n]] = E\left[(X[n] - \hat{X}[n])(X[n] - \hat{X}[n])^T\right] C_n^T = \epsilon^2[n]C_n^T.$$

Also

$$\begin{aligned} \sigma_{\tilde{Y}}^2[n] &= E[\tilde{Y}[n]\tilde{Y}^T[n]] \\ &= E\left[\left(C_n(X[n] - \hat{X}[n]) + V_n\right)\left(C_n(X[n] - \hat{X}[n]) + V_n\right)^T\right] \\ &= C_n\epsilon^2[n]C_n^T + \sigma_V^2[n]. \end{aligned}$$

So, $G[n] = \epsilon^2[n]C_n^T\left[C_n\epsilon^2[n]C_n^T + \sigma_V^2[n]\right]^{-1}$. We know,

$$\epsilon^2[n] = E\left[(X[n] - \hat{X}[n])(X[n] - \hat{X}[n])^T\right] = E\left[X[n]X^T[n]\right] - E\left[\hat{X}[n]\hat{X}^T[n]\right].$$

Now, $X[n] = A_nX[n-1] + B_nW[n]$ and $\hat{X}[n] = A_n\left[\hat{X}[n-1] + G_{n-1}\tilde{Y}[n-1]\right]$. Therefore,

$$E[X[n]X^T[n]] = A_nE\left[X[n-1]X^T[n-1]\right]A_n^T + B_n\sigma_W^2[n]B_n^T,$$

and also

$$E[\hat{X}[n]\hat{X}^T[n]] = A_nE\left[\hat{X}[n-1]\hat{X}^T[n-1]\right]A_n^T + A_nG_{n-1}E\left[\tilde{Y}[n-1]\tilde{Y}^T[n-1]\right]G_{n-1}^TA_n^T.$$

Hence,

$$\epsilon^2[n] = A_n\left(\epsilon^2[n-1] - G_{n-1}\sigma_{\tilde{Y}}^2[n-1]G_{n-1}^T\right)A_n^T + B_n\sigma_W^2[n]B_n^T.$$

From the result above,

$$G_{n-1}\sigma_{\tilde{Y}}[n-1]G_{n-1}^T = E\left[X[n-1]\tilde{Y}^T[n-1]\right]G_{n-1}^T = \epsilon^2[n-1]C_{n-1}^TG_{n-1}^T.$$

Thus, finally

$$\epsilon^2[n] = A_n\epsilon^2[n-1]\left(I - C_{n-1}^TG_{n-1}^T\right)A_n^T + B_n\sigma_W^2[n]B_n^T.$$

11. (a) Since $\tilde{Y}[-N] \perp \tilde{Y}[-N+1] \perp \ldots \tilde{Y}[N]$. Using Theorem 11.1-4 property (b), we have

$$\hat{E}\left[X[n]|\tilde{Y}[-N], \tilde{Y}[-N+1], \ldots, \tilde{Y}[N]\right] = \hat{E}[X[n]|\tilde{Y}[-N]] + \hat{E}\left[X[n]|\tilde{Y}[-N+1], \ldots, \tilde{Y}[N]\right].$$

By the same procedure

$$\hat{E}\left[X[n]|\tilde{Y}[-N], \tilde{Y}[-N+1], \ldots, \tilde{Y}[N]\right] = \sum_{k=-N}^{N} \hat{E}[X[n]|\tilde{Y}[k]].$$

Let $\hat{E}\left[X[n]|\tilde{Y}[k]\right] \triangleq g[k]\tilde{Y}[k]$. Then we have

$$\hat{E}\left[X[n]|\tilde{Y}[-N], \tilde{Y}[-N+1], \ldots, \tilde{Y}[N]\right] = \sum_{k=-N}^{N} g[k]\tilde{Y}[k].$$

(b) Let $\hat{X}[N] = \hat{E}\left[X[n]|\tilde{Y}[-N], \ldots, \tilde{Y}[N]\right]$.
$\because E\left[\hat{X}[N]|\hat{X}[0], \hat{X}[1], \ldots, \hat{X}[N-1]\right] = \hat{X}[N-1]$ since $E \triangleq \hat{E}$ in Gaussian case.
$\therefore \hat{E}\left[X[n]|\tilde{Y}[-N], \ldots, \tilde{Y}[N]\right]$ is a Martingale sequence. Using the result of problem 11.4, we conclude

$$\lim_{N\to\infty} \hat{E}\left[X[n]|\tilde{Y}[-N], \ldots, \tilde{Y}[N]\right] = \lim_{N\to\infty} \sum_{k=-N}^{N} g[k]\tilde{Y}[k]$$

exists with probability 1.

12. $\hat{R}_N[m] = \frac{1}{N}\sum_{n=0}^{N-1} x[n+m]x^*[n]$
(a)

$$E\left\{\hat{R}_N[m]\right\} = E\left\{\frac{1}{N}\sum_{n=0}^{N-1} x[n+m]x^*[n]\right\} = \frac{1}{N}\sum_{n=0}^{N-1} E\left\{x[n+m]x^*[n]\right\} = R_X[m]$$

(b) Show that $\lim_{N\to\infty} E\left\{|\hat{R}_N[m] - R_X[m]|^2\right\} = 0$.
$\lim_{N\to\infty} E\left\{|\hat{R}_N[m] - R_X[m]|^2\right\}$
$= \lim_{N\to\infty} E\left\{(\hat{R}_N[m] - R_X[m])(\hat{R}_N^*[m] - R_N^*[m])\right\}$
$= \lim_{N\to\infty} E\left\{\hat{R}_N[m]\hat{R}_N^*[m] - \hat{R}_N[m]\hat{R}_X^*[m] - \hat{R}_N^*[m]R_X[m] + R_X[m]R_X^*[m]\right\}$
$= \lim_{N\to\infty} E\left\{\hat{R}_N[m]\hat{R}_N^*[m] - R_X[m]R_X^*[m]\right\}$.
Now apply 4^{th} order moment property for (complex) Gaussian random variables to get
$E\left\{\hat{R}_N[m]\hat{R}_N^*[m]\right\}$
$= \frac{1}{N^2}E\left\{\sum_{n_1=0}^{N-1} X[n_1+m]X^*[n_1]\sum_{n_2=0}^{N-1} X^*[n_2+m]X[n_2]\right\}$
$= \frac{1}{N^2}\sum_{n_1,n_2} E\left\{X[n_1+m]X^*[n_1]X^*[n_2+m]X[n_2]\right\}$
$= \frac{1}{N^2}\sum_{n_1,n_2} E\left\{X[n_1+m]X^*[n_1]\right\}E\left\{X[n_2+m]X^*[n_2]\right\}$

$+\frac{1}{N^2}\sum_{n_1,n_2} E\{X[n_1+m]X^*[n_2+m]\}\, E\{X^*[n_1]X[n_2]\}$
$+$ extra term which is zero if $X[n]$ has symmetry condition as in (10.6-4) and (10.6-5) which can be restated for a complex random sequence $X[n]=X_r[n]+jX_i[n]$ as
$K_{X_rX_r}[m]=K_{X_iX_i}[m]$ and $K_{X_rX_i}[m]=-K_{X_iX_r}[m]$.
(NOTE: $K[\cdot]=R[\cdot]$ because $X[n]$ is zero mean.)
In this symmetric case, which occurs for the bandpass random process of section 10.6, we see that $E\{X[n_1+m]X[n_2]\}$ is zero as follows:
$E\{X[n_1+m]X[n_2]\}=$
$=E\{(X_r[n_1+m]+jX_i[n_1+m])(X_r[n_2]+jX_r[n_2])\}$
$=E\{X_r[n_1+m]X_r[n_2]-X_i[n_1+m]X_i[n_2]\}+jE\{X_i[n_1+m]X_r[n_2]+X_r[n_1+m]X_i[n_2]\}$
$=K_{X_rX_r}[n_1-n_2+m]-K_{X_iX_i}[n_1-n_2+m]+j\left(K_{X_iX_r}[n_1-n_2+m]+K_{X_rX_i}[n_1-n_2+m]\right)$
$=0+j0=0$
(For more on complex Gaussian random processes and random sequences, see *Discrete Random Signals and Statistical Signal Processing*, C. W. Therrien, Prentice-Hall, 1992.)

So,

$$
\begin{aligned}
E\left\{|\hat{R}_N[m]|^2\right\} &= R_X[m]R_X^*[m]+\frac{1}{N^2}\sum_{n_1,n_2=0}^{N-1} R_X[n_1-n_2]R_X^*[n_1-n_2] \\
&= |R_X[m]|^2+\sum_{n=-(N-1)}^{N-1}\frac{N-|n|}{N^2}|R_X[n]|^2.
\end{aligned}
$$

Thus $\lim_{N\to\infty} E\left\{|\hat{R}_N[m]-R_X[m]|^2\right\}\le \lim_{N\to\infty}\frac{1}{N}\sum_{m=-\infty}^{\infty}|R_X[m]|^2=0$ for square summable $R_X[m]$.
(NOTE: For real-valued random sequence, get two $O\left(\frac{1}{N}\right)$ terms.)

13. (a) The solution to this part is the same as the solution to Problem 8.32 (refer). The power spectral density is given by

$$
S_{XX}(w)=10\left(\frac{1-\rho_1^2}{1+\rho_1^2-2\rho_1\cos w}\right)+5\left(\frac{1-\rho_2^2}{1+\rho_2^2-2\rho_2\cos w}\right),
$$

where $\rho_1=e^{-\lambda_1},\rho_2=e^{-\lambda_2}$.

(b)

$$
\begin{aligned}
E\{I_N(w)\} &= \sum_{m=-(N-1)}^{N-1}\frac{N-|m|}{N}R_X[m]e^{-jwm} \\
&= \sum_{m=0}^{N-1}\frac{N-m}{N}\left(10e^{-\lambda_1 m-jwm}+5e^{-\lambda_2 m-jwm}\right)+\sum\left(10e^{\lambda_1 m-jwm}+5e^{\lambda_2 m-jwm}\right).
\end{aligned}
$$

So, $\lim_{N\to\infty}E\{I_N(w)\}=S_X(w)+\lim_{N\to\infty}\left(\sum_{m=0}^{N-1}\frac{-m}{N}\left(10e^{-\lambda_1 m-jwm}+5e^{-\lambda_2 m-jwm}\right)\right)$.
Now, $\left|\sum_{m=0}^{N-1}me^{-\lambda m-jwm}\right|\le\sum_{m=0}^{N-1}me^{-\lambda m}$ for any $\lambda>0$, and $\sum_{m=0}^{\infty}me^{-\lambda m}>\infty$.
Hence $\lim\frac{1}{N}\sum me^{-\lambda m}\to 0$. Thus we can conclude for this case that

$$
\lim_{N\to\infty}E\{I_N(w)\}=S_X(w).
$$

14. $r_0 a_1 = r_1$
$\sigma_X^2 a_1 = \sigma_X^2 \rho \implies a_1 \rho$
$\sigma_e^2 = r_0 - \sum_{m=1}^{p} a_m r_m = \sigma_X^2 - \rho^2 \sigma_X^2 = \sigma_X^2(1-\rho^2)$.

$$\begin{aligned} S_X(w) &= \frac{1}{\frac{1}{\sigma_e^2}\left|1-\sum_{m=1}^{p} a_m e^{-jwm}\right|^2} \\ &= \frac{\sigma_x^2(1-\rho^2)}{|1-\rho e^{-jw}|^2} \\ &= \frac{\sigma_x^2(1-\rho^2)}{1-2\rho\cos w+\rho^2}, |w| \le \pi. \end{aligned}$$

15. The MATLAB code (below) uses an AR(3) model to generate the N point random sample. Figure 1 is an estimate of the correlation function, computed with $N = 100$. The bottom axis should run from -100 to 100, as the zero shift value for $R[m]$ estimate is in the middle of the plot. Following the first plot, are three AR(3) spectral estimates for $N = 25, 100$, and 512 data points (Fig. 2). Also on each plot is the true psd for our AR(3) model.

```
%This program generates an AR3 estimate of psd of AR3 model.
clear
Pi=3.1415927;
disp('This .m file computes ar(3) parametric psd estimate.');
N=input('choose data length (<=512) = ');
randn('state',0);
w=randn(N,1);
bt=[1.0 0.0 0.0 0.0];
at=[1.0 -1.700 1.530 -0.648];
disp('The true a vector is a = [1.0 -1.7 1.53 -0.648].');
x=filter(bt,at,w);
y=flipud(x);
disp('Now calculating estimate of R[m].');
z=(1./N)*conv(x,y);
figure(1)
plot(z)
title('estimate of correlation function');
pause(5);
R(1,1)=z(N);
R(2,2)=z(N);
R(3,3)=z(N);
R(1,2)=z(N+1);
R(2,3)=z(N+1);
```

```
R(1,3)=z(N+2);
R(2,1)=R(1,2);
R(3,2)=R(2,3);
R(3,1)=R(1,3);
R
pause(5);
r(1,1)=z(N+1);
r(2,1)=z(N+2);
r(3,1)=z(N+3);
r
pause(5);
a=inv(R)*r
disp('May compare a vector values to ar(3) coefficients.');
b1=[1.0 0.0 0.0 0.0];
a1=[1.0 -a(1) -a(2) -a(3)];
[h,w]=freqz(b1,a1,512,'whole');
s=h.*conj(h);
figure(2)
plot(w,s);
title('ar(3) power spectral density estimate');
pause(3);
[ht,w]=freqz(bt,at,512,'whole');
st=ht.*conj(ht);
figure(3)
plot(w,st);
title('true ar(3) power spectral density');
pause(3);
figure(4)
plot(w,s,w,st)
title('ar(3) estimate and true ar(3) psd.');
text(0,-5,'Please press any key to exit.');
```

16. The Matlab function VITERBIPATH given below computes the most likely state sequences for the observations.

```
% The following function computes the most likely state seqeuence
% accounting for the observations.
% N is the number of states and must be an integer.
```

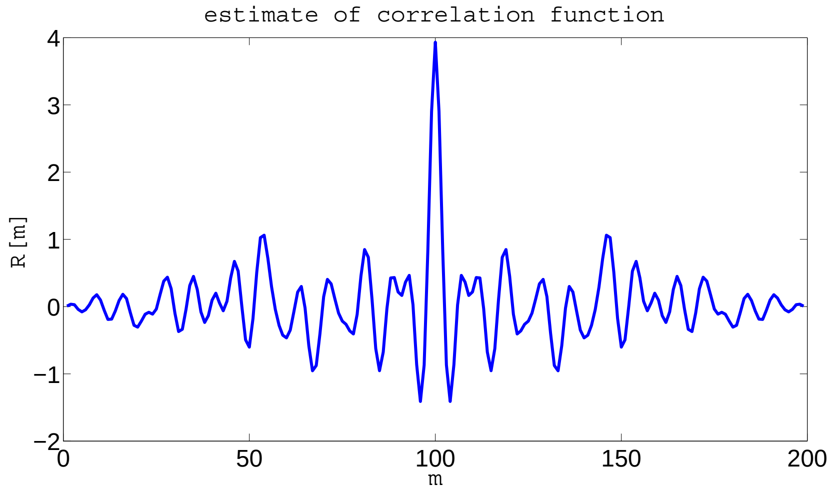


Figure 1: Estimate of correlation function for $N = 100$.

% L is either the number of observations or the maximum value f

% the discrete time index

% O is the observation vector and in this program must correspond to

% the columns of B. For example if you observe

% {heads, heads, tails, heads, tails} the observation vector would be

% {1,1,2,1,2} where a 1 would correspond to a head and a 2 would

% correspond to a tail.

% A is the sate transitionn probability matrix and is 2x2 in this case

% since there are only two states 'heads' and 'tails'.

% B is the state-conditional output probability matrix; the first column

% of B is [P[head|state1],P[head|state2]]

% PINT is the inital state probability row vector

function [PSTAR, Q] = VITERBIPATH(N,L,O,A,B,PINT)

% Initialize

Q = zeros(1,L);

psi = zeros(N,L);

phi = zeros(N,L);

for i = 1:N

phi(i,1) = PINT(i)*B(i,O(1));

end

% Iterate forward

for j = 2:L,

for i = 1:N,

```
for id = 1:N,
phiid(id) = phi(id,j-1) * A(id,i) * B(i,O(j));
psiid(id) = phi(id,j-1) * A(id,i);
end
phi(i,j) = max(phiid);
[junk, psi(i,j)] = max(psiid);
end
end
% Iterate backwards to find path
[PSTAR, Q(L)] = max(phi(:,L));
for n = L-1:-1:1
Q(n) = psi(Q(n+1), n+1);
end
% Graphical output
figure(1);
clf % clears current window
% The next instruction creates a graphical rectangular space of x-dimension
% going from 0 to L+1 and y-dimension from 0 to N+1.
axis([0 L+1 0 N+1]);
hold on; % keeps the current graphics
for j = 1:L
plot(j*ones(1,N),1:N,'o');
end
plot(1:L, Q,'-'); % Q is a vector whose components are the states
                  % at time l = 1, ..., L
% Hence a solid line segment drawn from the previous node to the new node
% at tile l.
set(gca, 'XTickLabel', [1:L]); % puts the numbers 1,2,...,L at the ticks
set(gca, 'YTick'    , [1:N]); % puts tick marks on y-label
set(gca, 'YTickLabel', [1:N]); % puts the numbers 1,2,...,N at the ticks
xlabel('time index');
ylabel('state'    );
title('most likely state path determined by Viterbi algorithm');
% grid
hold off; % liberates you from restrictions of
          % the graphical environment above
```

We run the MATLAB file to check the answer in Example 11.5-5.

>> VITERBIPATH(2, 3, [1,2,2], [0.6,0.4,0.3,0.7], [0.3,0.7,0.6,0.4], [0.7,0.3])

ans = 0.0370

This is the highest probability of observing this sequence probability. Now that we are confident that the program works, we can try it on the observation vector: head, head, head = 1,1,1. The output figure obtained is given in Fig. 3.

>> VITERBIPATH(2, 3, [1,1,1], [0.6,0.4,0.3,0.7], [0.3,0.7,0.6,0.4], [0.7,0.3])

ans = 0.0318

17. The most likely state path is given in Fig. 4.

>> VITERBIPATH(2, 3, [1,2,2,1,1], [0.6,0.4,0.3,0.7], [0.3,0.7,0.6,0.4], [0.7,0.3])

ans = 0.0318

18. We are given

$$Y_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2, Y_2 = \frac{1}{3}X_1 + \frac{2}{3}X_2,$$

or

$$X_1 = 2Y_1 - Y_2, X_2 = -Y_1 + 2Y_2,$$

where X_1: Poisson with λ_1, X_2: Poisson with λ_2, and $P_{X_1X_2}[m,k] = P_{X_1}[m]P_{X_2}[k]$.

We know that the maximum likelihood estimators for λ_1, λ_2 are given by

$$\hat{\lambda}_1^{(ml)} = X_1 = 2Y_1 - Y_2,$$

$$\hat{\lambda}_2^{(ml)} = X_2 = -Y_1 + 2Y_2.$$

Now let us see how the EM algorithm yields this result:

E-step: $\tilde{X}^{(k+1)} = E[\tilde{X}^k|\tilde{Y};\tilde{\lambda}^{(k)}]$,
M-step: $\tilde{\lambda}^{(k+1)} = \arg\max_{\tilde{\theta}}\left\{-\sum_{i=1}^n \theta_i + \tilde{X}^{(k+1)}\tilde{\Gamma}\right\}$,

where $\tilde{\Gamma} \triangleq (\log\theta_1, \log\theta_2, \ldots, \log\theta_n)$. To compute $E[\tilde{X}^{(h)}|\tilde{Y};\tilde{\lambda}^{(h)}]$, we need
$P\left[X_1^{(h)} = x_1, X_2^{(h)} = x_2|Y_1 = y_1, Y_2 = y_2; \lambda_1^{(h)}, \lambda_2^{(h)}\right]$
$= P\left[X_1^{(h)} = x_1|Y_1 = y_1; \lambda_1^{(h)}]P[X_2^{(h)} = x_2|Y_2 = y_2; \lambda_2^{(h)}\right]$;
since X_1, X_2 are independent, $X_1^{(h)}, X_2^{(h)}$ are also independent. However, in this case

$$P[X_1^{(h)} = x_1|\tilde{Y} = \tilde{y}] = \begin{cases} 1, & x_1 = 2y_1 - y_2 \\ 0, & \text{else.} \end{cases}.$$

Likewise

$$P[X_2^{(h)} = x_2|\tilde{Y} = \tilde{y}] = \begin{cases} 1, & x_2 = -y_1 + 2y_2 \\ 0, & \text{else.} \end{cases},$$

independent of the index h. Hence $X_1^{(h+1)} = X_1^{(h)}, X_2^{(h+1)} = X_2^{(h)}$ or for all k

$$X_1^{(k)} = 2Y_1 - Y_2, X_2^{(k)} = -Y_1 + 2Y_2.$$

Thus the E-step becomes

$$\tilde{X}^{(h+1)} = E[\tilde{X}^{(h)}|\tilde{Y}, \tilde{\theta}^{(h)}] = (2Y_1 - Y_2, -Y_1 + 2Y_2)^T .$$

And the M-step becomes

$$\tilde{\hat{\lambda}}^{(h+1)} = \arg\max_{\tilde{\theta}} \left\{ -\sum_{i=1}^{2} \theta_i + X_1^{(h)} \log\theta_1 + X_2^{(h)} \log\theta_2 \right\} \triangleq \arg\max_{\tilde{\theta}} \Lambda(\tilde{\theta}).$$

$\frac{\partial\Lambda}{\partial\theta_1} = 0 \implies 1 = \frac{X_1^{(h+1)}}{\theta_1}$ or $\theta_1 = \lambda_1^{(ml)} = 2Y_1 - Y_2$. $\frac{\partial\Lambda}{\partial\theta_2} = 0 \implies 1 = \frac{X_2^{(h+1)}}{\theta_2}$ or $\theta_2 = \lambda_2^{(ml)} = -Y_1 + 2Y_2$,
independent of k. Hence the E-M algorithm will converge after one iteration.

19. From the cited equation, we have

$$\sigma_U^2 \left| 1 - \sum_k a_k e^{-jwk} \right|^2 = \sigma_W^2 \left(1 - \sum_{k\neq 0} c_k e^{-jwk} \right)^2 ,$$

where σ_U^2 is the minimum M.S. interpolation error and σ_W^2 is the minimum M.S. prediction error, with the respective predictor coefficients $a_k, k = 1, \ldots, p$ and interpolator coefficients $c_k = c_{-k}, k = 1, \ldots, p$ ($c_0 = a_0 = 0$).

This equation is between two polynomials in the variable e^{-jw}. Setting the constant terms equal, we get $\sigma_k^2 \left(1 + \sum_{k=1}^{p} |a_k|^2\right) = \sigma_W^2$ or

$$\sigma_U^2 = \frac{\sigma_W^2}{1 + \sum |a_k|^2}$$

where $\sigma_U^2 < \sigma_W^2$ if any $a_k \neq 0, k = 1, \ldots, p$.

20. Set $D^{(k)}[n] \triangleq X[n] - X^{(k+1)}[n]$. Then from the two given equations,

$$D^{(k)}[n] = C * D^{(k-1)}[n] \frac{\sigma_v^2}{\sigma_u^2 + \sigma_v^2} = \sum_l c_l D^{(k-1)}[n-l] \frac{\sigma_v^2}{\sigma_u^2 + \sigma_v^2}.$$

Consider an interval $[-N, N]$ where the solution will be simulated for some large $N < \infty$. Then define the norm

$$||D^{(k)}|| \triangleq \max_{|n|\leq N} |D^{(k)}[n]|.$$

We have

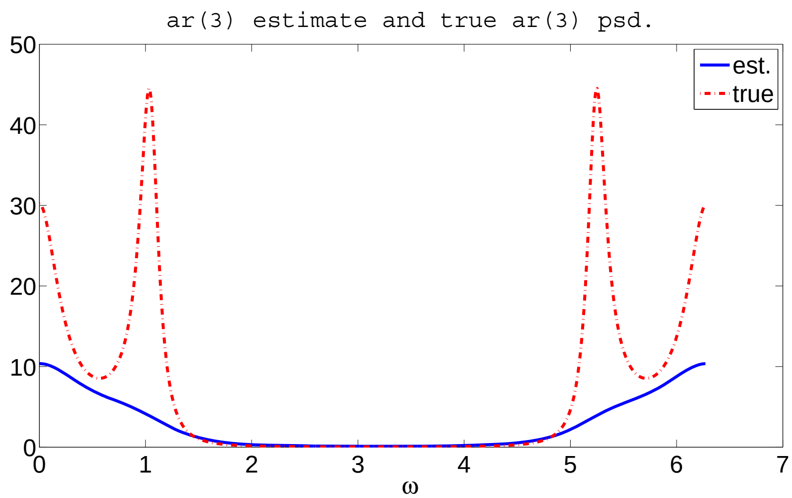
$$\begin{aligned} ||D^{(k)}|| &= \max_{|n|\leq N} \left| \sum_l c_l D^{(k-1)}[n-l] \right| \frac{\sigma_v^2}{\sigma_u^2 + \sigma_v^2} \\ &\leq \left(\sum |c_l| \right) \max_{|n|\leq N} \left| D^{(k-1)}[n-l] \right| \frac{\sigma_v^2}{\sigma_u^2 + \sigma_v^2} \\ &\leq \rho ||D^{(k-1)}||, \end{aligned}$$

where $\rho \triangleq \frac{\sigma_v^2}{\sigma_u^2 + \sigma_v^2} \sum_l |c_l| < 1$. Note that in the above equation, the maximum of $|n|$ should be less than or equal to $N - p$ where p is the order of the Markov model, but $N - p \approx N$.

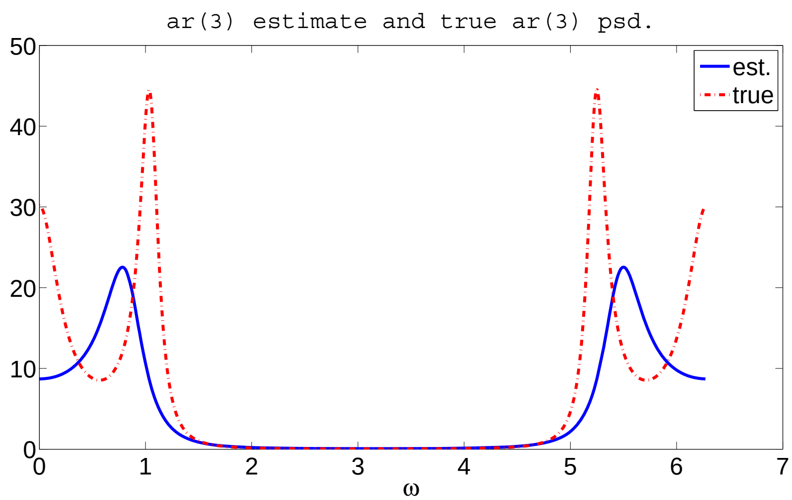
So, $||D^{(k)}|| \leq \rho ||D^{(k-1)}||, k \geq 1$. As $k \to \infty$, clearly $||D^{(k)}|| \to 0$ and so

$$X^{(k)}[n] \to X[n] \text{ for } -N < n < N.$$

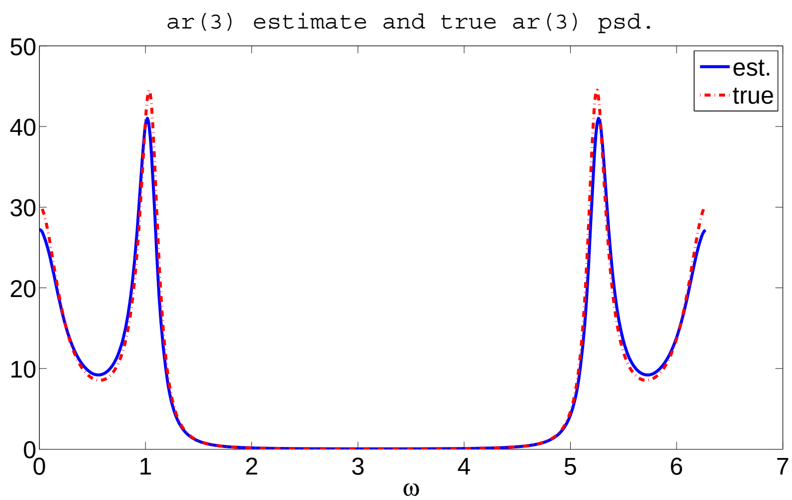
NOTE: We cannot let $N = +\infty$ because then the norm $||D||$ defined as $\max |D[n]|$ might be infinite. Still, N is large compared to the Markov order p should be sufficient. If we know the $Y[n]$ are finite valued, then we could let $N = +\infty$, because the equations are stable (BIBO stable).



(a) $N = 25$



(b) $N = 100$



(c) $N = 512$

Figure 2: Comparison of the true spectrum and AR(3) spectral estimate for different values of N.

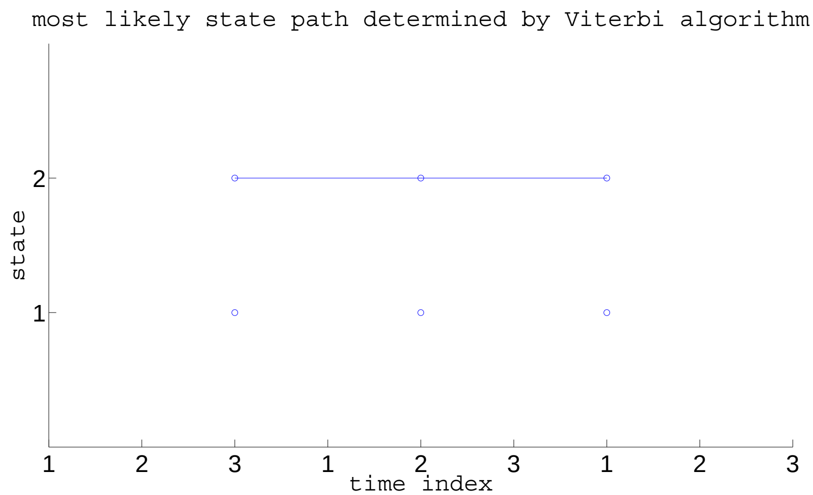


Figure 3:

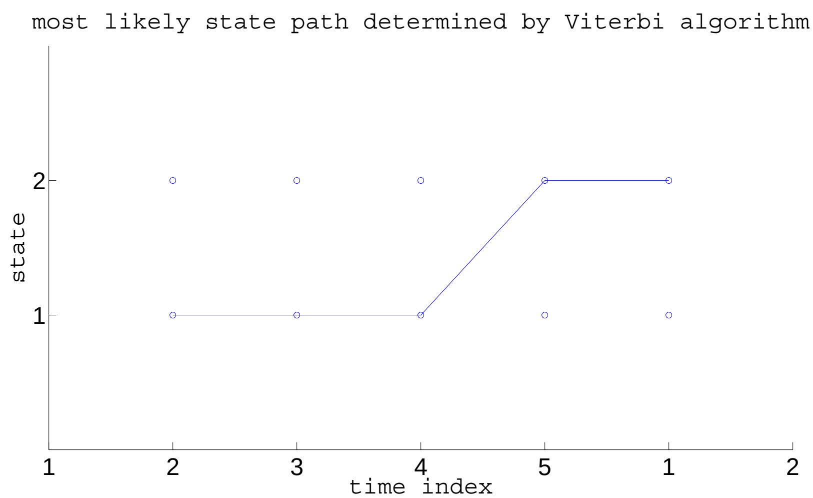


Figure 4: